METHYLATION AND SUMOYLATION REGULATE LSD1-COREST RECRUITMENT AND ACTIVITY TO CONTROL TRANSCRIPTIONAL AND CELLULAR FUNCTIONS OF GFI1

by

Matthew Edward Velinder

A dissertation submitted to the faculty of
The University of Utah
in partial fulfillment of the requirements for the degree of

Doctor of Philosophy

Department of Oncological Sciences

The University of Utah

August 2016

The University of Utah Graduate School

STATEMENT OF DISSERTATION APPROVAL

The dissertation of **Matthew Edward Velinder**

has been approved by the following supervisory committee members:

Michael Eugene Engel	, Chair	**6/1/2016** Date Approved
Bradley Cairns	, Member	**6/1/2016** Date Approved
Trudy Oliver	, Member	**6/1/2016** Date Approved
Timothy Formosa	, Member	**6/1/2016** Date Approved
Mahesh Chandrasekharan	, Member	**6/1/2016** Date Approved

and by **Bradley Cairns** , Chair of

the Department of **Oncological Sciences**

and by David B. Kieda, Dean of The Graduate School.

ABSTRACT

Proper cell fate decisions require precise and coordinated changes in gene expression. Alterations in gene expression proceed through the functions of transcription factors, their associated coregulators and histone modifying enzymes. However, how these complex and diverse groups of proteins and enzymes coordinate their respective functions at target genes remains largely unknown. To gain additional insights into the coordinated activities of transcription factors and histone modifying enzymes, we studied how the transcription factor growth factor independence 1 (GFI1) carries out transcriptional repression through interaction with coregulator histone modifying enzymes. GFI1 is a transcriptional repressor and master regulator of normal and malignant hematopoiesis. GFI1 is comprised of a transcriptionally repressive N-terminal Snail/Slug/GFI1 (SNAG) domain, a C-terminal concatemer of DNA binding zinc fingers and a linker region which separates them. The relatively simple protein domain structure of GFI1 makes it an ideal transcription factor for studying mechanisms of transcriptional repression. We describe here two novel mechanisms of transcriptional repression by GFI1, both of which occur through posttranslational modification. First, we identify and characterize a SUMOylation event carried out by the SUMO2/3, UBC9, and PIAS3 SUMOylation machinery. SUMOylation occurs at K239 within a type I SUMO consensus element in the linker region of GFI1. We find that SUMO

defective GFI1 derivatives fail to complement Gfi1 depletion phenotypes in zebrafish developmental erythropoiesis and in granulocyte differentiation in cultured human cells. SUMO defective GFI1 derivatives also display impaired LSD1/CoREST binding and fail to repress the GFI1 target gene *MYC* during granulocyte differentiation and enforced *MYC* expression blocks GFI1 mediated granulocyte differentiation. Second, we show SMYD2 mediated methylation at K8 within the GFI1 SNAG domain is a critical determinant of GFI1 transcriptional repression and contributes to GFI1 hematopoietic differentiation and leukemia cell survival functions. The methylation defective GFI1 SNAG domain lacks repressor function due to a failure of LSD1 recruitment and accumulation of promoter H3K4 dimethyl marks. Our findings here suggest GFI1 SUMOylation and methylation are part of a series of regulatory inputs that regulate GFI1 function. From these data we propose SNAG domain methylation and linker region SUMOylation coordinate LSD1/CoREST recruitment and enable CoREST dependent activation of LSD1 H3K4 demethylase activity for repression of GFI1 target genes. Our findings add GFI1 to the growing roster of transcription factors regulated by posttranslational modification and provides a rare mechanistic understanding into how these modifications regulate transcription factor functions through the recruitment histone modifying enzyme effectors.

To everyone who supported my scientific curiosity.

“This is water, this is water.”
David Foster Wallace

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF TABLES

ACKNOWLEDGEMENTS

First and foremost, I would like to thank my PhD mentor Dr. Michael Engel. Mike is a thoroughly genuine, caring and selfless person. He is a rigorous, thoughtful and demanding scientist who challenges those around him to do their best work possible, while maintaining unwavering support for the individual and their needs. Mike is a skilled, focused, and hypothesis driven scientist, a much needed and sadly often underappreciated role in the current preclinical research space. I am extremely grateful for the lasting personal and professional relationship I have developed with Mike during my time in his lab.

I am thankful to the members of my thesis committee, Dr. Brad Cairns, Dr. Trudy Oliver, Dr. Mahesh Chandrasekharan and Dr. Tim Formosa who have been generous with their thoughts and time throughout my graduate work. Brad Cairns is a critical and meticulous scientific thinker, always willing to help plan definitive and compelling experiments. Trudy Oliver is a driven, brilliant young investigator who, from a place of genuine caring and support, pushes those around her to do their very best. Mahesh Chandrasekharan is a knowledgeable and technically skilled bench scientist who is devoted to training and mentoring. Tim Formosa is a dedicated mentor, educator, and rigorous biochemist capable of providing valuable insights and thoughts from a wide breadth of knowledge. I would also like to thank

Dr. John Weis for the time he served on my thesis committee. John Weis was a thoroughly supportive mentor, generous with his time, and always speaking from a calm and caring place. I am also grateful to Dr. David Jones for the time I spent in his lab. David Jones provided me with time during my early graduate work to develop careful critical thinking skills as well as training in a number of technical skills including the use of zebrafish a model organism.

I cannot express the gratitude I feel for the love of my life, best friend, and (now) wife Sarah. She has shown unflinching support during my graduate work. Every day she reminds me of what really matters in life including, love, gratitude, and friendship. I am grateful for her endless kindness, patience, and encouragement during my thesis work. I look forward to all our future adventures together post PhD. I would also like to thank our two dogs Colby and Hela for their uncanny ability to lift my spirits.

I am extremely grateful to my family. They have been unconditionally supportive of my education. I thoroughly appreciate their emphasis on the importance of education and striving to reach one's goals. To my parents, Jay and Cheryl Velinder, my brother Jason, my sister Megan and her wonderful family, and my grandparents Curt and Sally Thompson, thank you for all you have contributed to my personal development. I cannot express how thankful I am to have been surrounded by such loving and supportive people throughout my life.

I am also thankful to my lab mates, friends and colleagues. I am grateful to have shared a common purpose with such thoughtful, supportive, and pleasant people.

CHAPTER 1

INTRODUCTION

Gene expression controls cellular identity and function

Transcription factors modulate gene expression to dictate cellular functions

Higher organisms, including humans, rely on the coordinated functions of distinct cell types to maintain organism homeostasis. Over 200 cell types are present within the human body, each with unique and indispensable functions that contribute to organism homeostasis. However, each of these functionally distinct cell types harbors an identical genome. How this remarkable level of functional diversity arises from an identical genome remains a profound question of modern molecular genetics. It is now appreciated that identical genomes achieve functional diversity by precisely and accurately regulating gene expression. Such a model of functional diversity through genetic simplicity predicts all cells types retain the potential to become any functionally distinct cell type. This prediction has now been realized in landmark findings demonstrating cell identity is remarkably plastic and cell identity can be altered by the introduction of a surprisingly small number of genetic factors (1).

Genetic factors with such profound impacts on cell identity and function are predominantly transcription factors, proteins capable of binding and regulating the transcriptional status of specific deoxyribonucleic acid (DNA) sequences. Of the roughly 20,000 protein coding genes in the human genome, it is estimated that 2,000 genes encode transcription factors (2). Many of these transcription factors have been appropriately dubbed “master regulators” and “pioneering” factors. One such master regulator was described roughly 30 years ago. In a series of elegant

papers, Weintraub and colleagues described a gene that was sufficient for converting fibroblasts to myoblasts (3, 4). How this gene accomplishes such a remarkable feat would later be clarified, but Weintraub's group named the gene myoblast determination gene number 1, or MyoD1, and speculated that, "Eventually, this information, together with input from local cues, could focus down to the activation of only one or a few specific genes that are ultimately responsible for determining a specific cell type." Master regulator transcription factors have been identified for many cell types, from myogenic cells (MyoD) back in 1986 to pluripotent cells (Oct4, Sox2, Klf4, Myc) in 2006 (1, 3, 4). Transcription factors are also often dysregulated in a variety of human diseases including cancer (5). These examples highlight the power of transcription factors to dictate cell identity, fate, and function.

Transcription factors exert their influence on cell fate and function through their ability to directly bind DNA and alter gene expression. Specific protein domains that confer DNA binding potential to proteins have been identified. These domains include, but are not limited to, C_2H_2 type zinc-nuclease fingers, homeodomains, helix loop helixes, basic leucine zippers, and nuclear hormone receptors (2). And while these domains are distinct in structure, they all share the ability to directly bind DNA in sequence specific manner, utilizing specific hydrogen bond donor and acceptor patterns within the major groove of consensus DNA to do so (2). Meta analysis of human transcription factors based on functional annotation reveals transcription factors contribute to a diverse range of biological processes, including, cellular, reproductive, developmental, metabolic, and

immune system processes as well as cellular localization and responses to stimuli (2). Furthermore, normal human tissues express roughly 150 to 300 transcription factors, with more functionally complex tissues such as fetal lung, brain, and placenta expressing more transcription factors, while more quiescent tissues such as appendix, skeletal muscle, and skin expressing fewer transcription factors (2). It is now well established that transcription factors coordinate expression of specific gene networks capable of altering and reinforcing cell fates and functions.

Chromatin structure regulates gene expression

Chromatin structure

Transcription factors were previously generalized as transcriptional activators or transcriptional repressors, either promoting or deterring transcription from nearby genes. And indeed, similar simplistic mechanisms of activator and repressor antagonism exist in bacteria and lower organisms. In eukaryotes, however, transcription factors exert their influence on gene expression through a variety of coregulators and histone modifying enzymes that alter chromatin structure. Chromatin is generally defined as the DNA of the host genome and the associated proteins that package the genome (6).

Chromatin is the functional unit of DNA in the genome and chromatin function is a reflection of its structure. Chromatin is comprised of nucleosomes, each of which is comprised of roughly 147 base pairs of DNA wrapped around histone protein octamer (Figure 1.1). Nucleosome histone octamers consist of two copies of each of the four canonical histone proteins: H2A, H2B, H3, and H4 which

form a globular protein scaffold for DNA to wrap around (Figure 1.1) (7). DNA maintains nucleosome contact through numerous weak electrostatic interactions between positively charged histone residues and negatively charged phosphate groups in the DNA backbone (6). Wrapping of DNA around a nucleosome constitutes the primary function of chromatin, which is the compaction of DNA. It is this packaging by histones and higher orders of chromatin compaction that allows roughly two linear meters of genomic DNA to be packaged into the microscopic nucleus of every cell, a level of compaction that is estimated to be over 10,000 fold.

From a level of increasing chromatin compaction, double stranded DNA is initially wrapped around individual nucleosomes with intervening DNA sequence not bound by nucleosomes. This wrapping constitutes a “beads on a string” structure, where nucleosomes are “beads” and intervening DNA is “string” (Figure 1.1). Condensing and repositioning nucleosomes within this “beads on a string” structure, as well as binding of histone H1 to “string” DNA results in formation of a 30 nanometer (nm) fiber (Figure 1.1). Additional rolling and looping of 30 nm fibers onto themselves results in extended and eventually condensed chromosome conformations (Figure 1.1). Finally, organization of these condensed looping chromatin formations leads to formation of the mitotic chromosomes consisting of shorter p- and longer q-arms connected by a highly condensed centromere.

Histone variants

To exert their biological functions, transcription factors must bind their specific consensus DNA sequence. However, high degrees of chromatin compaction deter transcription factor binding by concealing consensus DNA sequences on nucleosome surfaces or by burying consensus sequences deep within higher order chromatin structures called heterochromatin. Conversely, DNA sequences with less condensed euchromatin structures are considered accessible to transcription factors.

The composition of histone proteins within nucleosomes alter the relative composition of heterochromatin and euchromatin by augmenting chromatin compaction. Canonical nucleosomes within the genome are comprised of two copies of H2A, H2B, H3, and H4 histones. However, no fewer than 55 histone variants have been identified in humans (8). Incorporation of noncanonical histones within nucleosomes can dramatically alter chromatin function. For example, CENPA is a centromere specific histone H3 variant incorporated into nucleosomes at heterochromatic centromeres during mitosis (9). CENPA deposition at centromeres is required for kinetochore and spindle assembly formation and experimentally mistargeting CENPA to noncentromeric regions assembles centromere kinetochore machinery at these sites (10). CENPA is also deposited at sites of DNA damage (11). CENPA deposition is independent of γ H2AX deposition, a phosphorylated histone variant also deposited at sites of DNA damage (12). These examples demonstrate the profound impact on chromatin function histone variants can have.

DNA methylation and demethylation

While DNA packaging histones are critical for chromatin functions, modification of DNA itself can also influence chromatin function. DNA methylation is the covalent addition of methyl groups to the fifth carbon position of cytosine bases within DNA. DNA methylation predominantly occurs in a CpG dinucleotide context where a cytosine base is followed by a guanine base. This context allows for the maintenance of DNA methylation marks following DNA replication, a maintenance function carried out by the DNA methyltransferase enzyme DNMT1 (13). DNMT1 specifically recognizes and methylates the unmethylated cytosine base within a hemimethylated CpG, a CpG where only one complementary strand cytosine is methylated. DNA methylation marks can also be deposited at unmethylated CpGs de novo by the DNA methyltransferases DNMT3A and DNMT3B (13). DNA methylation is transcriptionally repressive, owing to its ability to recruit chromatin compacting enzymes and cofactors. DNA methylation is critical for a variety of developmental processes, including most notably X chromosome inactivation (14).

Long considered a permanent mark, it is now understood that DNA methylation marks can be removed. Early reports correctly asserted that passive DNA demethylation may occur from a failure to maintain methylation following DNA replication. However, it is now understood that active DNA demethylation also occurs and is critical for a variety of cellular and developmental processes (15). While alternative mechanisms have been proposed, the preponderance of evidence suggests DNA demethylation occurs through the oxidation of 5-

methylcytosine bases followed by base excision repair. Oxidation of 5-methylcytosine (5mC) bases occurs through the activity of ten-eleven translocase (TET) enzymes, resulting in the formation of 5-hydroxymethylcytosine (5hmC) bases. 5hmC bases can also be further oxidized to generate 5-formylcytosine (5fC) and 5-carboxylcytosine (5caC) bases (15). 5fC and 5caC bases are recognized and removed by the base excision repair activity of thymine DNA glycosylase (TDG) enzymes, generating an abasic site. Repair of the abasic site using an unmethylated cytosine base then results in demethylation (15).

Histone posttranslational modification and the histone code hypothesis

Similar to DNA methylation, covalent modification of histone proteins dynamically regulates chromatin structure and function. Pioneering work by Vincent Allfrey in 1964 demonstrated histones are subject to acetylation and methylation and these modifications are associated with specific transcriptional states (16). Canonical and noncanonical histone proteins have been found to be modified by a variety of posttranslational modifications, including acetylation, methylation, phosphorylation, ubiquitination, SUMOylation, and ribosylation (8). These modifications can occur within the globular domains of histones, however; the functional consequences of modifications within the N-terminal tail of histone proteins has been more thoroughly described. Histone tails extend from the globular nucleosome and are largely unstructured, allowing access for enzymes responsible for writing, reading, and erasing posttranslational modifications. In

humans, an increasingly complex suite of over 150 enzymes regulate histone posttranslational marks (8). Furthermore, more than 100 unique histone residues are modified by posttranslational modification (8). Cataloguing the enzymes responsible for the deposition and removal of methyl and acetyl marks on a subset of H3 lysines, K4, K9, K27, and K36 highlights the impressive amount of combinatorial diversity present within the H3 tail (Figure 1.2). These lysine methylation and acetylation events are a small subset of the posttranslational modifications that occur on histone proteins. Descriptions of where these modifications occur in the genome, the transcriptional status of nearby genes and whether these modifications occur independently or coincident with other modifications has led to the “histone code” hypothesis. This hypothesis posits that the distinct combinations of histone modifications, either directly, or indirectly through the action of other effector proteins, dictate chromatin function, including the transcriptional status of associated genes.

The logic of the histone code hypothesis has classified proteins as histone writers, readers, and erasers (17). Writers are enzymes capable of catalytically adding posttranslational modifications to histone substrates at specific residues. Readers are proteins capable of binding specific histone modifications, while erasers are enzymes capable of removing histone modifications (17).

Of the impressive breadth of histone posttranslational modifications, lysine methylation is perhaps the most well characterized histone modification. Lysine methylation occurs on H1, H2B, H3, and H4 histones (8). Histone H3 tail methylation occurs at K4, K9, K20, K27, and K36 (Figure 1.2). Which lysine residue

is methylated as well as the stoichiometry (mono, di or tri) of methylation at that lysine impact the transcriptional status of nearby genes and the functions of nearby regulatory DNA elements. For example, H3K4 monomethylation, H3K4me1, marks enhancer regulatory sequences. However, dimethylation and trimethylation of H3K4, H3K4me2, and H3K4me3 is associated with transcriptionally active promoters (18). Trimethylation at lysine 9 of H3, H3K9me3, is associated transcriptionally silent genes, highlighting the importance of the residue on which trimethylation occurs (19). Specific histone modifications occur concurrently at active and repressed gene promoters, consistent with the histone code hypothesis (Figure 1.3). So called “bivalent” histone signatures have also been described. These bivalent domains tend to be marked by activating H3K4me3 marks and repressive H3K27me3 marks and are considered “poised” for expression at specific developmental or differentiation time points. One recently described example is self renewing Thy1+ adult germline stem cells where the Sox2 developmental gene is marked with H3K4me3 and repressive H3K27me3 marks as well as DNA methylation (20). These marks maintain Sox2 in a bivalent poised state, primed for activation, yet transcriptionally silent. How these bivalent regions are maintained and resolved remains an area of active inquiry.

Despite extensive characterization of posttranslational modifications that occur on histones, how these marks work together and in isolation to coordinate transcriptional outcomes remains incompletely understood. One current hypothesis posits single histone modifications act as individual permission slips for specific transcriptional outcomes, with multiple permission slips being required for

outcomes to occur. This principle is observed for transcriptionally activating H3K4 methylation carried out by the MLL family of histone methyltransferases. The presence of transcriptionally repressive H3K9 methyl marks on nearby histones deters MLL activity at H3K4. However, if H3K9 is acetylated, an activating mark, MLL activity is enhanced.

Histone modifications carry out transcriptional responses through their ability to be read by histone mark readers. Histone readers have been identified for the majority of known histone marks (8). H3K4me3 marks, for instance, are read by a variety of protein domains, including chromo, PHD, and tudor domains (21). The ING family of proteins contain PHD domains and function within larger histone acetyltransferase complexes (21, 22). Additionally, CHD1, a transcriptionally activating ATP dependent chromatin remodeler, contains a chromodomain, which preferentially binds H3K4me3 marked histones (23). CHD1 also interacts with basal transcription and preinitiation complex machinery to promote transcription of H3K4me3 marked genes (24). Similarly, acetylated histones are recognized by bromodomain containing proteins including TATA binding protein associated factors, histone acetyltransferases, and general transcription factors (25). This example of H3K4 trimethylation highlights how coherent histone marks promote transcriptional outcomes.

Posttranslational modification of transcription factors

Modes and mechanisms of transcription factor posttranslational modification

It is important to note that while histone mark readers are profoundly impactful on transcriptional outcomes, these readers lack sequence specific DNA binding capability and instead use histone marks to bind target regions in the genome. Furthermore, the enzymes that deposit histone marks, histone writers, also lack sequence specific DNA binding function. Therefore, how histone marks are precisely deposited and read at specific sequences within the genome remains an open question. Transcription factors are an obvious solution to this question, as they are capable of recognizing and binding specific DNA sequences. In fact, multiple reports have demonstrated transcription factors interact with chromatin modifiers at target genes (26). One such example is the epithelial to mesenchymal transition (EMT) transcription factor SNAIL. SNAIL directly interacts with the H3K9 methyltransferase polycomb repressor complex proteins EZH2 and SUZ12 at the CDH1/E-cadherin gene promoter and these interactions are required for proper CDH1 repression during EMT (27). These reports highlight how transcription factors facilitate chromatin modifier recruitment to target genes, however, how transcription factors carry out these recruitment events remains largely unknown.

Posttranslational modifications dynamically regulate protein functions, often by altering protein-protein interactions. As such, posttranslational modification is an appealing mechanism for how transcription factors interact with chromatin modifiers. Posttranslational modification of transcription factors would allow for

dynamic regulation of transcriptional states through regulated recruitment of specific chromatin modifiers. Indeed, an increasing number of transcription factors and other nonhistone proteins have been found to be posttranslationally modified (28-32).

In fact, the number of nonhistone proteins subject to one such posttranslational modification, methylation, now far outnumber that of histone proteins. Curiously, nonhistone protein methylation appears to occur predominantly through the activity of methyltransferases originally identified for their histone methyltransferase activity. The H3K4 methyltransferase SETD7 methylates at least 30 nonhistone proteins, 17 of which are transcription factors (30, 32). Similarly, the H3K9 methyltransferase G9a methylates at least 17 nonhistone proteins, 7 of which are transcription factors (30). Transcription factors methylated at lysines include: p53, RB, E2F1, TAF7, TAF10, PCAF, YAP, STAT3, FOXO3, GATA4, MTA1, C/EBPβ, MYOD, ERα, AR, KLF12, NFκB, GLI3, and RARα (Figure 1.4 and Table 1.1) (30-33). Many transcription factors within this list are master regulators of cell fate, differentiation and survival, highlighting the importance of lysine methylation in transcription factor function.

Examples of transcription factors regulated by lysine methylation

One of the most well characterized examples of transcription factor lysine methylation is p53. p53 is methylated at lysines 370, 372, 373, and 382 by the methyltransferases SMYD2, SETD7, G9a, and SETD8, respectively (Table 1.1)

(34-36). SMYD2 mediated methylation at K370 corresponds with decreased p53 DNA binding, decreased association with p53 effector 53BP1, and decreased p53 dependent transcriptional and DNA damage responses (34, 37). Similarly, methylation at K373 is associated with inactive p53 (38). K370 methylation is antagonized by SETD7 mediated methylation at K372, which promotes p53 DNA binding and p53 mediated cell death responses (39). Antagonistic methylation at K370 and K372 is dynamically regulated by lysine specific demethylase 1, LSD1, which removes K370 methylation marks (40).

Another master regulator transcription factor whose function is regulated by lysine methylation is RB. RB is methylated at K860 by SMYD2, which promotes RB interaction with the polycomb repressor complex protein L3MBTL1 (41). Similarly, methylation of RB at K873 by SETD7 promotes RB interaction with the heterochromatin protein HP1 (42).

The examples of p53 and RB demonstrate discrete methylation events can alter protein-protein and transcriptional coregulator interactions for carrying out specific transcriptional and biological outcomes. At present, a relatively small subset of methyltransferases and demethylases have been found to regulate nonhistone protein methylation status. Predominantly, these methyltransferases are: SMYD2, SMYD3, SETD7, SETD8, G9a, Glp, and NSD1, while LSD1 is the only described nonhistone protein demethylase (Figure 1.4 and Table 1.1) (30-32). These findings suggest transcription factor posttranslational modification and specifically lysine methylation is a means to specifically recruit transcriptional coregulators and chromatin modifiers.

GFI1 as a paradigm for studying mechanisms of transcriptional control

The GFI1 protein family

Growth factor independence 1 (GFI1) is a repressive transcription factor with many important functions in normal and malignant hematopoiesis (43). GFI1 was originally identified through an insertional Moloney murine leukemia virus (MMLV) screen where cloning of a provirus insertion site led to identification of the novel cDNA encoding a zinc finger containing protein, subsequently named GFI1 (44). Proviral insertion resulting in GFI1 expression conferred IL-2 independent growth to the previously IL-2 dependent T-cell lymphoma cell lines 2780d and 5675d (44).

GFI1 is comprised of an N-terminal Snail/Slug/GFI1 (SNAG) domain, a C-terminal concatemer of six C_2H_2-type zinc fingers and a linker region which separates them (Figure 1.5). GFI1 is the founding member of the GFI1 family, with the sole other member being GFI1B (45). GFI1 and GFI1B are highly conserved, with their SNAG domains sharing 95% amino acid identity (Figure 1.5) (46). The zinc fingers of GFI1 and GFI1B are also highly conserved, sharing 89% amino acid conservation (Figure 1.5). As such, both proteins bind a consensus DNA recognition motif with an AATC core recognition sequence (taAATCac(t/a)gca (43). GFI1 and GFI1B differ predominantly within their linker regions, which share only 39% conservation (Figure 1.5).

Due to the large degree of conservation in their SNAG domains and zinc fingers, GFI1 and GFI1B exert similar, yet distinct molecular functions, due at least

in part to their cell type and tissue specific expression patterns (Figure 1.6) (47). GFI1B is expressed and functions predominantly in hematopoietic stem and multipotent progenitor cells as well as in megakaryocyte-erythrocyte progenitor cells (Figure 1.6) (43, 45). GFI1 is expressed and functions predominantly in hematopoietic stem cells, myeloid, and lymphoid primed multipotent progenitor cells (Figure 1.6) (43).

The distinct hematologic functions of GFI1 and GFI1B are manifested physiologically in humans as well. Hereditary neutropenia due to GFI1 mutations has been reported in at least two families (48). Affected individuals are germline carriers of either an N382S or K403R mutation occurring within the fifth and sixth zinc fingers of GFI1, respectively. It is presumed, although remains to be functionally tested, that both mutations are dominant loss of function mutations owing to their ability to abolish DNA binding (N382S) or by interfering with cofactor binding (K403R) (43). It is also presumed the reported R421X, L400F, and P107A mutations give rise to neutropenia due to their ability to abolish DNA binding. One case of a neutropenic patient carrying a Q17 mutation within the SNAG domain has also been found (personal communication). To our knowledge, this is the only known mutation occurring in the SNAG domain that is associated with congenital neutropenia in humans. An S36N GFI1 mutation is observed in acute myeloid leukemia where it is thought to predispose patients through disrupting GFI1 coregulator interactions (49, 50). Gfi1 null mice recapitulate the neutropenic phenotype observed in humans, in addition to displaying defects in pre-T-cell differentiation and impaired inner ear hair cell development (47, 51). These results

highlight the importance of Gfi1 in hematopoiesis and specifically within the lymphoid and myeloid compartments.

Mutations in GFI1B have also been described clinically. A zinc finger 5 protein truncating frameshift mutation of GFI1B that is unable to bind DNA is observed in a family with an inherited mild thrombocytopenia and red blood cell anisopoikilocytosis (43). A protein truncating dominant negative mutation occurring at residue Q287 within zinc finger 5 of GFI1B has also been described in a family with gray platelet syndrome. This mutation abolishes DNA binding by GFI1B, causing platelets to have decreased alpha granules and confers a moderate bleeding tendency and a moderate thrombocytopenia in carriers (52). Gfi1b null mice are embryonic lethal due to failure to produce enucleated erythrocytes (53). The fetal liver of Gfi1b null embryos also contains developmentally arrested erythroid and megakaryocytic precursors, highlighting the importance of Gfi1b in erythrocyte and megakaryocyte cell lineage fate determination (53). Taken together the Gfi1 and Gfi1b human and mice phenotypes also demonstrate the distinct functions of GFI1 family members both inside and outside the hematopoietic compartment.

GFI1 transcriptional coregulators

Clinical descriptions of loss of function mutations within the zinc fingers of GFI1 highlight the critical importance of DNA binding for GFI1 function. However, aside from the requirement for DNA binding, relatively little is known about how GFI1 functions as a transcriptional repressor. A variety of transcriptional

corepressors and histone modifying enzymes interact with GFI1 (Figure 1.5). Histone deacetylase complexes 1, 2, and 3 (HDAC 1-3) interact with the N-terminal portion of the GFI linker region. Euchromatic histone-lysine N-methyltransferase 2 (G9a/EHMT2) binds GFI1 via contributions from both the N- and C- terminal portions of the GFI linker (54). Ajuba, an HDAC dependent transcriptional corepressor binds GFI1 through yet to be mapped regions of GFI1 (55). The activated Notch1-intracellular domain (N1-ICD) binds GFI1 within the C-terminal portion of the GFI1 linker (data unpublished). A variety of other transcription factors and transcriptional corepressors interact with GFI1 via the GFI1 linker and zinc fingers, including, MTG8, MTG16, ETS1, Miz1, and Pu.1 (Figure 1.5) (45, 56-58). Lysine specific demethylase 1 (LSD1/KDM1A) is a transcriptionally repressive mono and di H3K4 demethylase which binds GFI1 via the SNAG domain. LSD1 is the only protein known to bind the SNAG domain (Figure 1.5) (43).

GFI1 requires the SNAG domain and zinc fingers
for repression

Initial characterization of GFI1 demonstrated at least two domains are required for GFI1 transcriptional repressor function, the SNAG domain, and zinc fingers. DNA binding via zinc fingers 3-5 is required for transcriptional repression by GFI1 (59). Furthermore, each zinc finger (3, 4, and 5) is required for DNA binding, as deletion of a single one of these zinc fingers abolishes repressor activity (59). In addition to zinc fingers, the N-terminal SNAG domain is required for transcriptional repression by GFI1 (59). The linker region of GFI1 is largely

dispensable for repressor function, as a linker deleted construct consisting of only the SNAG domain and zinc fingers 1-6 has comparable repression ability as full length wild-type GFI1 (59). Additionally, zinc fingers 1-6 alone are not capable of transcriptional repression (59). These data suggest the SNAG domain is the dominant repressive element within GFI1, and is localized to sites of repression through the function of the GFI1 zinc fingers.

Individual residues within the SNAG domain are critical for GFI1 repressor function. A proline to alanine substitution at the 2 position (P2A) completely abolishes repressor function (59). Interestingly, alanine substitution at other SNAG domain positions does not impact repressor function, even at immediately adjacent positions, such as R3, S4, K10, and K11 (59). These findings demonstrate the SNAG domain is a critical determinant of GFI1 mediated transcriptional repression and suggests individual residues within the SNAG domain are required for this function.

Putative GFI1 SNAG domain methylation

Lysine 8 (K8) is a universally conserved residue across GFI1 family members as well as across human, mouse, fish, xenopus, dolphin, pig, dog, cat, horse, elephant, and guinea pig, among others (Figure 1.5). K8 resides within a KSKK motif embedded within the SNAG domain, occurring at residues 8 through 11 (Figure 1.5). As described previously, p53 is subject to methylation, which dynamically regulates its function (36). Lysine methylation occurs at residues K370, K372, and K373 of p53, which reside within a KSKK motif identical to that

found within the GFI1 SNAG domain (35). Within this methylation dependent functional axis of p53, LSD1 specifically demethylates the first lysine, K370, of the p53 KSKK motif (40). LSD1 is also a known effector of GFI1 family mediated gene repression and the only protein known to bind the GFI1 SNAG domain (43, 46). From this, we speculated the KSKK motif embedded within the GFI1 SNAG domain may be a mechanism by which LSD1 binds the GFI1 SNAG domain. We focused on K8 within this motif, as LSD1 acts on the K8 equivalent residue, K370, in p53 and as alanine substitution at K10 and K11 does not impair repressor function (59). Additionally, alanine substitution at K8 abolished LSD1 interaction with this GFI1B isoform (60). LSD1 from UT-7 cell lysates also preferentially bound a K8 dimethylated Gfi1B SNAG domain peptide *in vitro*, suggesting methylation at K8 may be a critical determinant of LSD1 recruitment to the SNAG domain (60).

GFI1 SUMOylation/ubiquitination

Similar to methylation, ubiquitin, and small ubiquitin related modifier (SUMO) marks are a posttranslational modification covalently attached to target substrates to modulate their function (61, 62). Previous reports indicate GFI1 interacts with the ubiquitin ligase Triad1 (63). However, the GFI1/Triad1 interaction was not mapped on either protein, nor was the site of ubiquitination identified, and counterintuitively the authors described Triad1 as prolonging cellular GFI1 half-life (63). Previous to our work, no evidence for GFI1 SUMOylation has been reported.

Diverse cellular functions regulated by SUMOylation include, protein stability, subcellular and subnuclear localization, chromatin structure and

transcriptional regulation (62). SUMOylation of target protein substrates proceeds through coordinated activities of three SUMO machinery enzymes (61, 62). An E1 SUMO activating enzyme coordinates transfer of the SUMO group to UBC9, an E2 SUMO conjugating enzyme. In the final step, an E3 SUMO ligase coordinates E2 and target protein specificity for the eventual transfer of the SUMO group to the target substrate (62). SUMO modification occurs at the ε-amine group of lysine residues and approximately 75% of SUMO modifications occur within a type I SUMO consensus element ψKx(D/E), where ψ is a large hydrophobic residue and x is any amino acid (64, 65).

An increasing number of proteins, including transcription factors have been found to be regulated by SUMOylation (61, 62). Sharp-1 is a myogenic transcription factor that antagonizes MyoD activity (66). Sharp-1 is SUMOylated at K240 and K255, and these modifications are enhanced by the presence of PIAS3 and PIASxα (66). SUMOylation defective Sharp-1 is impaired in transcriptional repression as well as MyoD antagonism and displays decreased binding to G9a, a methyltransferase that methylates K104 of MyoD (30, 66). These data demonstrate a key role for SUMOylation of Sharp-1 for G9a recruitment and inhibition of myogenic differentiation, presumably through MyoD. SOX10 is another transcription factor modified by SUMOylation. SUMOylation modulates SOX10 function through impairing interaction with SOX10 coactivators EGR2 and PAX3, thereby impairing SOX10 transcriptional activity (67). Similar to SOX10, SUMOylation of SOX6 impairs SOX6 transcriptional activity (68). These data demonstrate SUMOylation profoundly impacts transcription factor function, often

through augmenting interaction with transcriptional coregulators.

We analyzed the primary sequence of GFI1 using SUMO prediction algorithms and found a type I SUMO consensus element (GVKVES) occurring at residues 237-242 in the GFI1 linker, where the underlined sequence is the core type I consensus sequence (Figure 1.5) (65, 69). Furthermore, PIAS3, an E3 SUMO ligase binds the linker region of GFI1, where the type I SUMO consensus element centered around K239 occurs. (70). Given these published data, we hypothesized GFI1 may be subject to regulation by SUMOylation, likely within the linker region.

Posttranslational modifications may regulate GFI1 functions

During my thesis work I investigated how posttranslational modifications regulate GFI1 transcriptional repressor and cellular functions. Based on the previously published observations described above we focused on how putative SNAG domain methylation and linker region SUMOylation events modulate GFI1 function. We find that GFI1 is regulated by methylation and SUMOylation and these posttranslational modifications regulate GFI1 function through augmenting LSD1/CoREST interaction. In Chapter 2, we show GFI1 is SUMOylated at K239 by SUMO2/3, UBC9, and PIAS3 SUMO machinery. In Chapter 3, we show GFI1 is methylated at K8 by SMYD2. We find these posttranslational modifications are required for GFI1 transcriptional repressor function and contribute to GFI1 functions during primitive erythropoiesis in zebrafish as well as in GFI1 mediated

neutrophil differentiation and leukemia cell survival. Disruption of these modifications also impairs LSD1/CoREST binding, demonstrating the importance of LSD1/CoREST within this functional GFI1 axis. We propose SUMOylation and methylation are part of a dual input mechanism that regulates GFI1 function. These posttranslational modification dependent recruitment events then stabilize and activate LSD1 for H3K4 demethylation and repression of GFI1 target genes.

The context, scope, and implications of these findings are discussed in Chapter 4. This includes the potential combinatorial regulation of lysine methylation within the GFI1 SNAG domain, as well as potential crosstalk between methylation and SUMOylation in GFI1. We also speculate about functional crosstalk between GFI1, p53, and other nonhistone proteins regulated by lysine methylation. Lastly, we assess the potential breadth of transcription factor posttranslational modifications in the human proteome and the applications and limitations of current methodologies to characterize these modifications and their functional consequences. Taken together, the work of my thesis provides novel mechanisms for how posttranslational modifications of GFI1 serve to repress GFI1 target genes through LSD1/CoREST. My thesis work also contributes to the emerging understanding of how transcription factors interact with histone modifying enzymes and how these interactions are regulated through posttranslational modifications.

References

1. **Takahashi K, Yamanaka S.** 2006. Induction of pluripotent stem cells from mouse embryonic and adult fibroblast cultures by defined factors. Cell **126:**663–676.

2. **Vaquerizas JM, Kummerfeld SK, Teichmann SA, Luscombe NM.** 2009. A census of human transcription factors: function, expression and evolution. Nat Rev Genet **10:**252–263.

3. **Lassar AB, Paterson BM, Weintraub H.** 1986. Transfection of a DNA locus that mediates the conversion of 10T12 fibroblasts to myoblasts. Cell **47:**649.

4. **Davis RL, Weintraub H, Lassar AB.** 1987. Expression of a single transfected cDNA converts fibroblasts to myoblasts. Cell **51:**987.

5. **Bhagwat AS, Vakoc CR.** 2015. Targeting Transcription Factors in Cancer. Trends Cancer **1:**53–65.

6. **Khorasanizadeh S.** 2004. The nucleosome: from genomic organization to genomic regulation. Cell **116:**259–272.

7. **Clapier CR, Cairns BR.** 2009. The biology of chromatin remodeling complexes. Annu Rev Biochem **78:**273–304.

8. **Khare SP, Habib F, Sharma R, Gadewal N, Gupta S, Galande S.** 2011. HIstome--a relational knowledgebase of human histone proteins and histone modifying enzymes. Nucleic Acids Res **40:**D337–42.

9. **Sullivan KF, Hechenberger M, Masri K.** 1994. Human CENP-A contains a histone H3 related histone fold domain that is required for targeting to the centromere. J Cell Biol **127:**581–592.

10. **Van Hooser AA, Ouspenski II, Gregson HC, Starr DA, Yen TJ, Goldberg ML, Yokomori K, Earnshaw WC, Sullivan KF, Brinkley BR.** 2001. Specification of kinetochore-forming chromatin by the histone H3 variant CENP-A. J Cell Sci **114:**3529–3542.

11. **Zeitlin SG, Baker NM, Chapados BR, Soutoglou E, Wang JYJ, Berns MW, Cleveland DW.** 2009. Double-strand DNA breaks recruit the centromeric histone CENP-A. Proc Natl Acad Sci USA **106:**15762–15767.

12. **Rogakou EP, Pilch DR, Orr AH, Ivanova VS, Bonner WM.** 1998. DNA double-stranded breaks induce histone H2AX phosphorylation on serine 139. J Biol Chem **273:**5858–5868.

13. **Smith ZD, Meissner A.** 2013. DNA methylation: roles in mammalian development. Nat Rev Genet **14:**204–220.

14. **Sharp AJ, Stathaki E, Migliavacca E, Brahmachary M, Montgomery SB, Dupre Y, Antonarakis SE.** 2011. DNA methylation profiles of human active and inactive X chromosomes. Genome Research **21:**1592–1600.

15. **Kohli RM, Zhang Y.** 2013. TET enzymes, TDG and the dynamics of DNA demethylation. Nature **502:**472–479.

16. **Allfrey VG, Faulkner R, Mirsky AE.** 1964. Acetylation and methylation of histones and their possible role in the regulation of RNA synthesis. Proc Natl Acad Sci USA **51:**786–794.

17. **Cairns BR.** 2009. The logic of chromatin architecture and remodelling at promoters. Nature **461:**193–198.

18. **Bannister AJ, Kouzarides T.** 2011. Regulation of chromatin by histone modifications. Cell Research **21:**381–395.

19. **Hon GC, Hawkins RD, Ren B.** 2009. Predictive chromatin signatures in the mammalian genome. Hum Mol Genet **18:**R195–201.

20. **Hammoud SS, Yi C, Guccione E.** 2014. Chromatin and Transcription Transitions of Mammalian Adult Germline Stem Cells and Spermatogenesis. Cell Stem Cell **15:**239.

21. **Yun M, Wu J, Workman JL, Li B.** 2011. Readers of histone modifications. Cell Research **21:**564–578.

22. **Zhang T, Cooper S, Brockdorff N.** 2015. The interplay of histone modifications - writers that read. EMBO Rep **16:**1467–1481.

23. **Venters BJ, Pugh BF.** 2009. How eukaryotic genes are transcribed. Crit Rev Biochem Mol Biol **44:**117–141.

24. **Lin JJ, Lehmann LW, Bonora G, Sridharan R, Vashisht AA, Tran N, Plath K, Wohlschlegel JA, Carey M.** 2011. Mediator coordinates PIC assembly with recruitment of CHD1. Genes & Development **25:**2198–2209.

25. **Filippakopoulos P, Knapp S.** 2014. Targeting bromodomains: epigenetic readers of lysine acetylation. Nat Rev Drug Discov **13:**337–356.

26. **Tam WL, Weinberg RA.** 2013. The epigenetics of epithelial-mesenchymal plasticity in cancer. Nat Medicine **19:**1438–1449.

27. **Herranz N1, Pasini D, Díaz VM, Francí C, Gutierrez A, Dave N, Escrivà M, Hernandez-Muñoz I, Di Croce L, Helin K, García de Herreros A, Peiró S.** 2008. Polycomb Complex 2 Is Required for E-cadherin Repression by the Snail1 Transcription Factor. Mol Cell Bio **28:**4772.

28. **Benayoun BA, Veitia RA.** 2009. A post-translational modification code for transcription factors: sorting through a sea of signals. Trends Cell Biol **19:**189–197.

29. **Johnston SJ, Carroll JS.** 2015. Transcription factors and chromatin proteins as therapeutic targets in cancer. Biochim Biophys Acta **1855:**183–192.

30. **Biggar KK, Li SS-C.** 2014. Non-histone protein methylation as a regulator of cellular signalling and function. Nat Rev Mol Cell Biol **16:**5–17.

31. **Hamamoto R, Saloura V, Nakamura Y.** 2015. Critical roles of non-histone protein lysine methylation in human tumorigenesis. Nat Rev Cancer **15:**110–124.

32. **Stark GR, Wang Y, Lu T.** 2010. Lysine methylation of promoter-bound transcription factors and relevance to cancer. Cell Research **21:**375–380.

33. **Fu L, Wu H, Cheng SY, Gao D, Zhang L, Zhao Y.** 2016. Set7 mediated Gli3 methylation plays a positive role in the activation of Sonic Hedgehog pathway in mammals. Elife **5**.

34. **West LE, Gozani O.** 2011. Regulation of p53 function by lysine methylation. Epigenomics **3:**361–369.

35. **Chuikov S, Kurash JK, Wilson JR, Xiao B, Justin N, Ivanov GS, McKinney K, Tempst P, Prives C, Gamblin SJ, Barlev NA, Reinberg D.** 2004. Regulation of p53 activity through lysine methylation. Nature **432:**353–360.

36. **Scoumanne A, Chen X.** 2008. Protein methylation: a new mechanism of p53 tumor suppressor regulation. Histol Histopathol **23:**1143–1149.

37. **Huang J, Perez-Burgos L, Placek BJ, Sengupta R, Richter M, Dorsey JA, Kubicek S, Opravil S, Jenuwein T, Berger SL.** 2006. Repression of p53 activity by Smyd2-mediated methylation. Nature **444:**629–632.

38. **Huang J, Dorsey J, Chuikov S, Perez-Burgos L, Zhang X, Jenuwein T, Reinberg D, Berger SL.** 2010. G9a and Glp methylate lysine 373 in the tumor suppressor p53. J Biol Chem **285:**9636–9641.

39. **Kurash JK, Lei H, Shen Q, Marston WL, Granda BW, Fan H, Wall D, Li E, Gaudet F.** 2008. Methylation of p53 by Set7/9 mediates p53 acetylation and activity in vivo. Molecular Cell **29:**392–400.

40. **Huang J, Sengupta R, Espejo AB, Lee MG, Dorsey JA, Richter M, Opravil S, Shiekhattar R, Bedford MT, Jenuwein T, Berger SL.** 2007. p53 is regulated by the lysine demethylase LSD1. Nature **449:**105–108.

41. **Saddic LA, West LE, Aslanian A, Yates JR, Rubin SM, Gozani O, Sage J.** 2010. Methylation of the retinoblastoma tumor suppressor by SMYD2. J Biol Chem **285:**37733–37740.

42. **Munro S, Khaire N, Inche A, Carr S, La Thangue NB.** 2010. Lysine methylation regulates the pRb tumour suppressor protein. Oncogene **29:**2357–2367.

43. **Moroy T, Vassen L, Wilkes B, Khandanpour C.** 2015. From cytopenia to leukemia: the role of Gfi1 and Gfi1b in blood formation. Blood **126:**2561.

44. **Gilks CB, Bear SE, Grimes HL, Tsichlis PN.** 1993. Progression of interleukin-2 (IL-2)-dependent rat T cell lymphoma lines to IL-2-independent growth following activation of a gene (Gfi-1) encoding a novel zinc finger protein. Mol Cell Bio **13:**1759–1768.

45. **van der Meer LT, Jansen JH, van der Reijden BA.** 2010. Gfi1 and Gfi1b: key regulators of hematopoiesis. Leukemia **24:**1834–1843.

46. **Saleque S, Kim J, Rooke HM, Orkin SH.** 2007. Epigenetic regulation of hematopoietic differentiation by Gfi-1 and Gfi-1b is mediated by the cofactors CoREST and LSD1. Molecular Cell **27:**562–572.

47. **Fiolka K, Hertzano R, Vassen L, Zeng H, Hermesh O, Avraham KB, Dührsen U, Möröy T.** 2006. Gfi1 and Gfi1b act equivalently in haematopoiesis, but have distinct, non-overlapping functions in inner ear development. EMBO Rep **7:**326–333.

48. **Person RE, Li F-Q, Duan Z, Benson KF, Wechsler J, Papadaki HA, Eliopoulos G, Kaufman C, Bertolone SJ, Nakamoto B, Papayannopoulou T, Grimes HL, Horwitz M.** 2003. Mutations in proto-oncogene GFI1 cause human neutropenia and target ELA2. Nat Genet **34:**308–312.

49. **Khandanpour C, Krongold J, Schütte J, Bouwman F, Vassen L, Gaudreau M-C, Chen R, Calero-Nieto FJ, Diamanti E, Hannah R, Meyer SE, Grimes HL, van der Reijden BA, Jansen JH, Patel CV, Peeters JK, Löwenberg B, Dührsen U, Göttgens B, Möröy T.** 2012.

The human GFI136N variant induces epigenetic changes at the Hoxa9 locus and accelerates K-RAS driven myeloproliferative disorder in mice. Blood **120:**4006–4017.

50. **Khandanpour C, Thiede C, Valk PJM, Sharif-Askari E, Nückel H, Lohmann D, Horsthemke B, Siffert W, Neubauer A, Grzeschik K-H, Bloomfield CD, Marcucci G, Maharry K, Slovak ML, van der Reijden BA, Jansen JH, Schackert HK, Afshar K, Schnittger S, Peeters JK, Kroschinsky F, Ehninger G, Löwenberg B, Dührsen U, Möröy T.** 2010. A variant allele of Growth Factor Independence 1 (GFI1) is associated with acute myeloid leukemia. Blood **115:**2462–2472.

51. **Karsunky H, Zeng H, Schmidt T, Zevnik B, Kluge R, Schmid KW, Dührsen U, Möröy T.** 2002. Inflammatory reactions and severe neutropenia in mice lacking the transcriptional repressor Gfi1. Nat Genet **30:**295–300.

52. **Monteferrario D, Bolar NA, Marneth AE, Hebeda KM, Bergevoet SM, Veenstra H, Gorkom BAPL-V, MacKenzie MA, Khandanpour C, Botezatu L, Fransen E, Van Camp G, Duijnhouwer AL, Salemink S, Willemsen B, Huls G, Preijers F, Van Heerde W, Jansen JH, Kempers MJE, Loeys BL, Van Laer L, van der Reijden BA.** 2013. A dominant-negative GFI1B mutation in the gray platelet syndrome. N Engl J Med **370:**245–253.

53. **Saleque S, Cameron S, Orkin SH.** 2002. The zinc-finger proto-oncogene Gfi-1b is essential for development of the erythroid and megakaryocytic lineages. Genes & Development **16:**301–306.

54. **Duan Z, Zarebski A, Montoya-Durango D, Grimes HL, Horwitz M.** 2005. Gfi1 coordinates epigenetic repression of p21Cip/WAF1 by recruitment of histone lysine methyltransferase G9a and histone deacetylase 1. Mol Cell Bio **25:**10338–10351.

55. **Montoya-Durango DE, Velu CS, Kazanjian A, Rojas MEB, Jay CM, Longmore GD, Grimes HL.** 2008. Ajuba functions as a histone deacetylase-dependent co-repressor for autoregulation of the growth factor-independent-1 transcription factor. J Biol Chem **283:**32056–32065.

56. **McGhee L, Bryan J, Elliott L, Grimes HL, Kazanjian A, Davis JN, Meyers S.** 2003. Gfi-1 attaches to the nuclear matrix, associates with ETO (MTG8) and histone deacetylase proteins, and represses transcription using a TSA-sensitive mechanism. Journal of Cellular Biochemistry **89:**1005–1018.

57. **Nakazawa Y, Suzuki M, Manabe N, Yamada T, Kihara-Negishi F, Sakurai T, Tenen DG, Iwama A, Mochizuki M, Oikawa T.** 2007. Cooperative interaction between ETS1 and GFI1 transcription factors in the repression of Bax gene expression. Oncogene **26:**3541–3550.

58. **Liu Q, Basu S, Qiu Y, Tang F, Dong F.** 2010. A role of Miz-1 in Gfi-1-mediated transcriptional repression of CDKN1A. Oncogene **29:**2843–2852.

59. **Grimes HL, Chan TO, Zweidler-McKay PA, Tong B, Tsichlis PN.** 1996. The Gfi-1 proto-oncoprotein contains a novel transcriptional repressor domain, SNAG, and inhibits G1 arrest induced by interleukin-2 withdrawal. Mol Cell Bio **16:**6263.

60. **Laurent B, Randrianarison-Huetz V, Frisan E, Andrieu-Soler C, Soler E, Fontenay M, Dusanter-Fourt I, Duménil D.** 2012. A short Gfi-1B isoform controls erythroid differentiation by recruiting the LSD1-CoREST complex through the dimethylation of its SNAG domain. J Cell Science **125:**993–1002.

61. **Gareau JR, Lima CD.** 2010. The SUMO pathway: emerging mechanisms that shape specificity, conjugation and recognition. Nat Rev Mol Cell Biol **11:**861–871.

62. **Wilkinson KA, Henley JM.** 2010. Mechanisms, regulation and consequences of protein SUMOylation. Biochem J **428:**133–145.

63. **Marteijn JAF, van der Meer LT, van Emst L, van Reijmersdal S, Wissink W, de Witte T, Jansen JH, van der Reijden BA.** 2007. Gfi1 ubiquitination and proteasomal degradation is inhibited by the ubiquitin ligase Triad1. Blood **110:**3128–3135.

64. **Nakakido M, Deng Z, Suzuki T, Dohmae N, Nakamura Y, Hamamoto R.** 2015. Dysregulation of AKT Pathway by SMYD2-Mediated Lysine Methylation on PTEN. Neoplasia **17:**367–373.

65. **Xu J, He Y, Qiang B, Yuan J, Peng X, Pan X-M.** 2008. A novel method for high accuracy sumoylation site prediction from protein sequences. BMC Bioinformatics **9:**8.

66. **Wang Y, Shankar SR, Kher D, Ling BMT, Taneja R.** 2013. Sumoylation of the basic helix-loop-helix transcription factor sharp-1 regulates recruitment of the histone methyltransferase G9a and function in myogenesis. J Biol Chem **288:**17654–17662.

67. **Girard M, Goossens M.** 2006. Sumoylation of the SOX10 transcription factor regulates its transcriptional activity. FEBS Lett **580:**1635–1641.

68. **Fernández-Lloris R, Osses N, Jaffray E, Shen LN, Vaughan OA, Girwood D, Bartrons R, Rosa JL, Hay RT, Ventura F.** 2006. Repression of SOX6 transcriptional activity by SUMO modification. FEBS Lett **580:**1215–1221.

69. **Xue Y, Zhou F, Fu C, Xu Y, Yao X.** 2006. SUMOsp: a web server for sumoylation site prediction. Nucleic Acids Res **34:**W254–7.

70. **Rodel B.** 2000. The zinc finger protein Gfi-1 can enhance STAT3 signaling by interacting with the STAT3 inhibitor PIAS3. The EMBO Journal **19:**5845.

Figure 1.1 Genomic DNA is subject to extensive packaging via nucleosomes. Levels of DNA compaction are illustrated. In order from least compact to most compact: DNA double helix, beads on a string, 30 nanometer fiber, extended chromosome, condensed chromosome and mitotic chromosome.

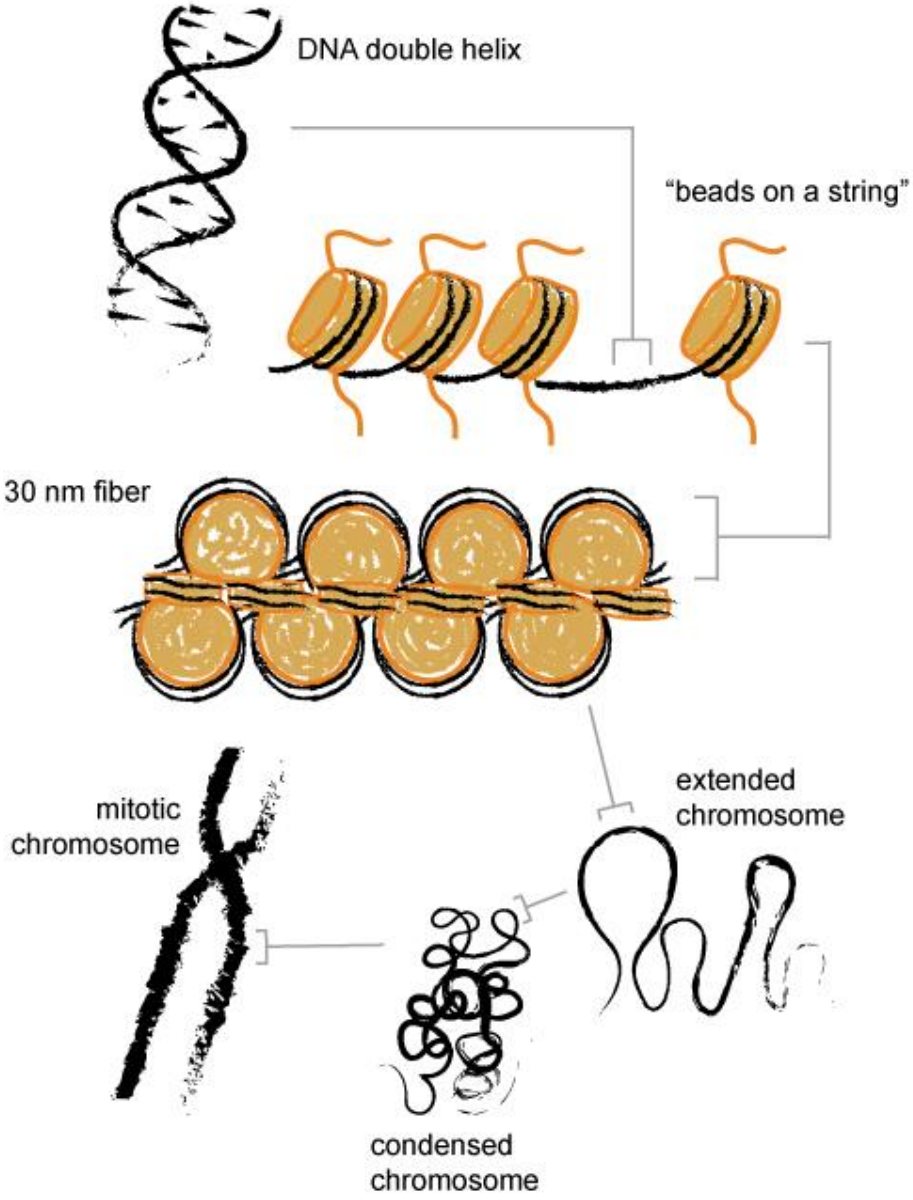

DNA double helix
"beads on a string"
30 nm fiber
mitotic chromosome
extended chromosome
condensed chromosome

Figure 1.2. A subset of lysine methylation and acetylation events occurring on lysines 4, 9, 27, and 36 of the histone H3 tail and the enzymes that write and erase these marks. The sequence of amino acids 1-40 for HIST1H3A is given. Methylation and demethylation events at lysine residues are depicted by dark red and light red boxes respectively. Acetylation and deacetylation events at lysine residues are depicted by dark blue and light blue boxes respectively. Enzymes with respective methylation/demethylation and acetylation/deacetylation activities are listed extended from their respective colored boxes. Specific mono (me1), di (me2) and tri (me3) methylation, demethylation and acetylation/deacetylation activities of enzymes are shown in parentheses.

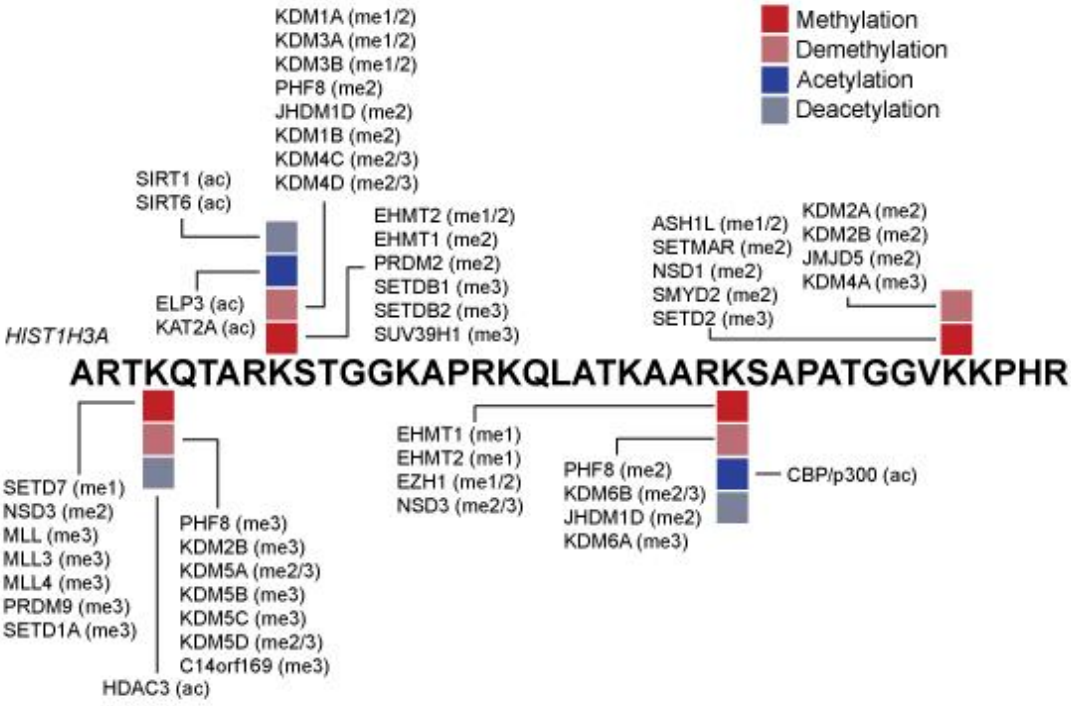

Methylation
Demethylation
Acetylation
Deacetylation
KDM1A (me1/2)
KDM3A (me1/2)
KDM3B (me1/2)
PHF8 (me2)
JHDM1D (me2)
KDM1B (me2)
KDM4C (me2/3)
KDM4D (me2/3)
SIRT1 (ac)
SIRT6 (ac)
EHMT2 (me1/2)
EHMT1 (me2)
PRDM2 (me2)
SETDB1 (me3)
SETDB2 (me3)
SUV39H1 (me3)
ELP3 (ac)
KAT2A (ac)
ASH1L (me1/2)
SETMAR (me2)
NSD1 (me2)
SMYD2 (me2)
SETD2 (me3)
KDM2A (me2)
KDM2B (me2)
JMJD5 (me2)
KDM4A (me3)
HIST1H3A
ARTKQTARKSTGGKAPRKQLATKAARKSAPATGGVKKPHR
SETD7 (me1)
NSD3 (me2)
MLL (me3)
MLL3 (me3)
MLL4 (me3)
PRDM9 (me3)
SETD1A (me3)
PHF8 (me3)
KDM2B (me3)
KDM5A (me2/3)
KDM5B (me3)
KDM5C (me3)
KDM5D (me2/3)
C14orf169 (me3)
HDAC3 (ac)
EHMT1 (me1)
EHMT2 (me1)
EZH1 (me1/2)
NSD3 (me2/3)
PHF8 (me2)
KDM6B (me2/3)
JHDM1D (me2)
KDM6A (me3)
CBP/p300 (ac)

Figure 1.3. Histone code promoter marks frequently associated with transcriptional activation and repression outcomes. Histone marks are listed in boxes above area in which they are typically found relative to the gene transcription start site. The transcription start site is depicted by a line with a pointed arrow (for active transcription) or a flat line (for repressed transcription). Promoter marks associated with active transcription are listed in the top schematic while promoter marks associated with repressed transcription are listed in the bottom schematic.

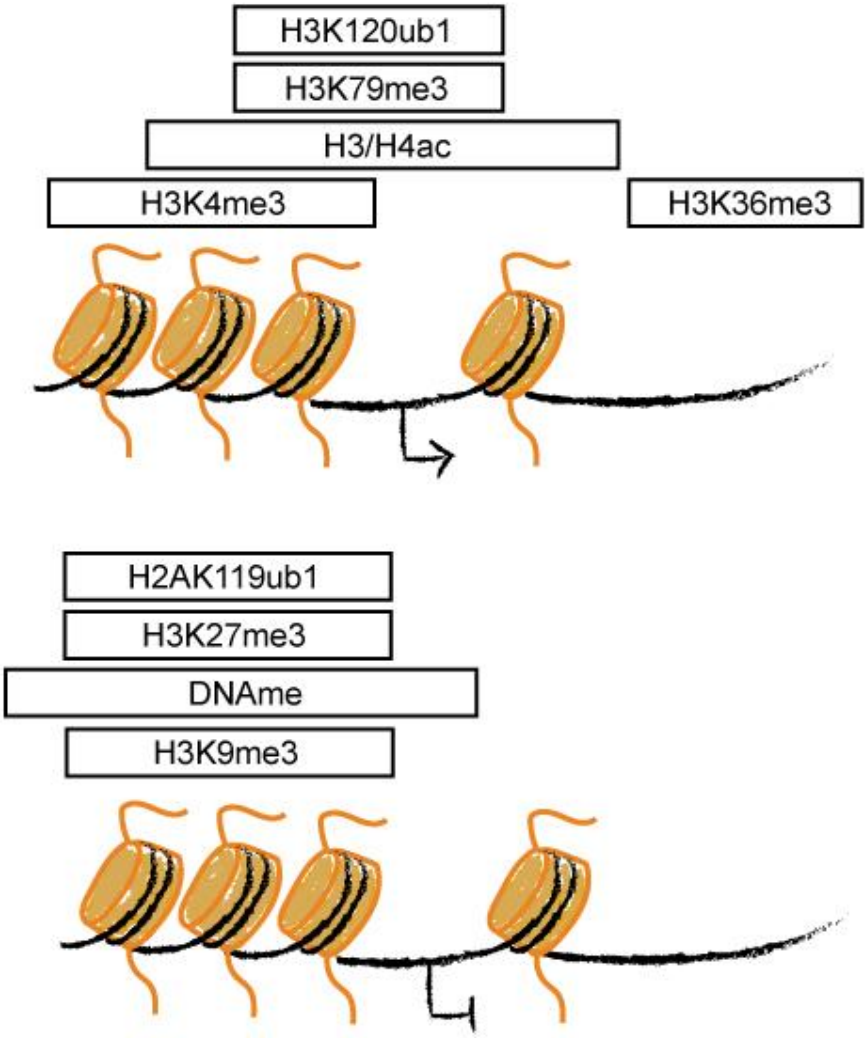

H3K120ub1
H3K79me3
H3/H4ac
H3K4me3
H3K36me3
H2AK119ub1
H3K27me3
DNAme
H3K9me3

Figure 1.4. Graphical representation of nonhistone proteins regulated by lysine methylation and the methyltransferases and demethylases that regulate methylation status of these proteins. Methylated proteins are shown as grey boxes with protein names above the boxes. GFI1 and p53 shown as black boxes. Methyltransferases and demethylases are shown as circles. LSD1 is the only demethylase shown, as it is the only demethylase known to regulate nonhistone protein methylation status. Lines connecting boxes and circles indicate the respective methyltransferase or demethylase regulates methylation status of the connected protein.

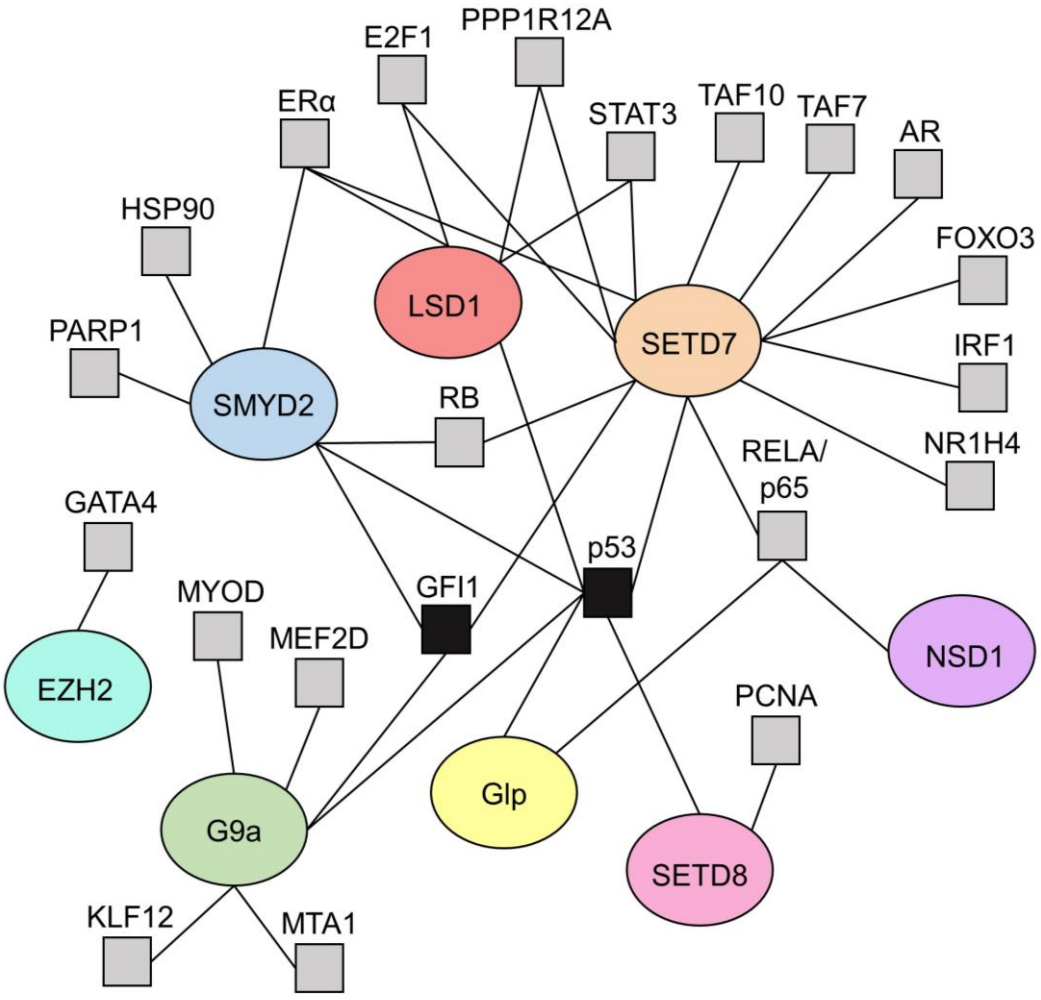
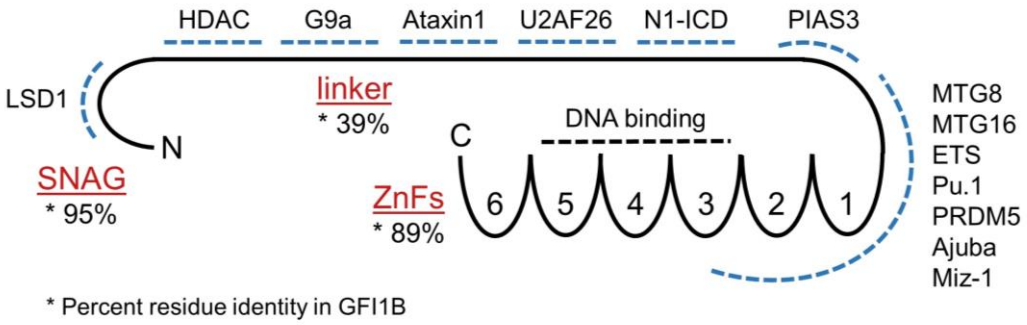

E2F1
PPP1R12A
ERα
STAT3
TAF10
TAF7
AR
HSP90
FOXO3
LSD1
SETD7
PARP1
IRF1
SMYD2
RB
NR1H4
RELA/
p65
GATA4
p53
MYOD
GFI1
NSD1
MEF2D
EZH2
PCNA
Glp
G9a
SETD8
KLF12
MTA1

Figure 1.5. Graphical depiction of GFI1 protein domains and where GFI1 interacting proteins bind within GFI1. GFI1 is comprised of an N-terminal SNAG domain, followed by a linker region and a C-terminal concatemer of six zinc fingers. Zinc fingers 3-5 are required for DNA binding. Various transcription factors, histone modifying enzymes and transcriptional coregulators and the regions within GFI1 they bind are highlighted. * percentages below GFI1 domains denote percent residue identity between GFI1 and GFI1B. Primary protein sequence of the GFI1 SNAG domain across species is given. A grey box designates a species conserved lysine-serine-lysine-lysine (KSKK) motif occurring at residues 8-11 in the GFI1 SNAG domain. Residues surrounding a type I SUMO consensus element (GVKVES) within the GFI1 linker region is given.



* Percent residue identity in GFI1B

SNAG domain:

		8	11	
Human	1 MPRSFLV	K	SKK	AHSYHQPRS 20
Mouse	MPRSFLV	K	SKK	AHSYHQPRS
Z.Fish	MPRSFLV	K	SKK	AHSYHQPRS
Xenopus	MPRSFLV	K	SKK	AHSYHQPRV
G.Pig	MPRSFLV	K	SKK	AHSYHQPRS
Dolphin	MPRSFLV	K	SKK	AHSYHQPRS
Pig	MPRSFLV	K	SKK	AHSYHQPRS
Dog	MPRSFLV	K	SKK	AHSYHQPRS
Cat	MPRSFLV	K	SKK	AHSYHQPRS
Ferret	MPRSFLV	K	SKK	AHSYHQPRS
Elephant	MPRSFLV	K	SKK	AHSYHQPRS
Pika	MPRSFLV	K	SKK	AHSYHQPRS
Horse	MPRSFLV	K	SKK	AHSYHQPRS

Linker region:

^{237}GVKVES^{242}

^{228}GLHADKGAGVKVESELLCTRLL^{249}

Figure 1.6. GFI1 and GFI1B contribute to distinct hematopoietic lineages. Respective colored areas for GFI1 and GFI1B indicate relative protein expression within the associated hematopoietic compartment as well as the functional contribution of GFI1 and GFI1B to this cell type and fate. Cell types are indicated by colored circles with the relative potential to contribute to each lineage indicated as a colored portion of the full circle. Cell type abbreviations are as follows. HSC – hematopoietic stem cell, LMPP – lymphoid primed multipotent progenitor, CLP – common lymphoid progenitor cell, EPLM – early progenitor with lymphoid and myeloid potential, B – B cell, DN1 – double negative 1, DN2 – double negative 2, T – T cell, Gran – granulocyte, CMkE – common megakaryocyte erythrocyte progenitor, Megak – megakaryocyte, Eryth – erythrocyte

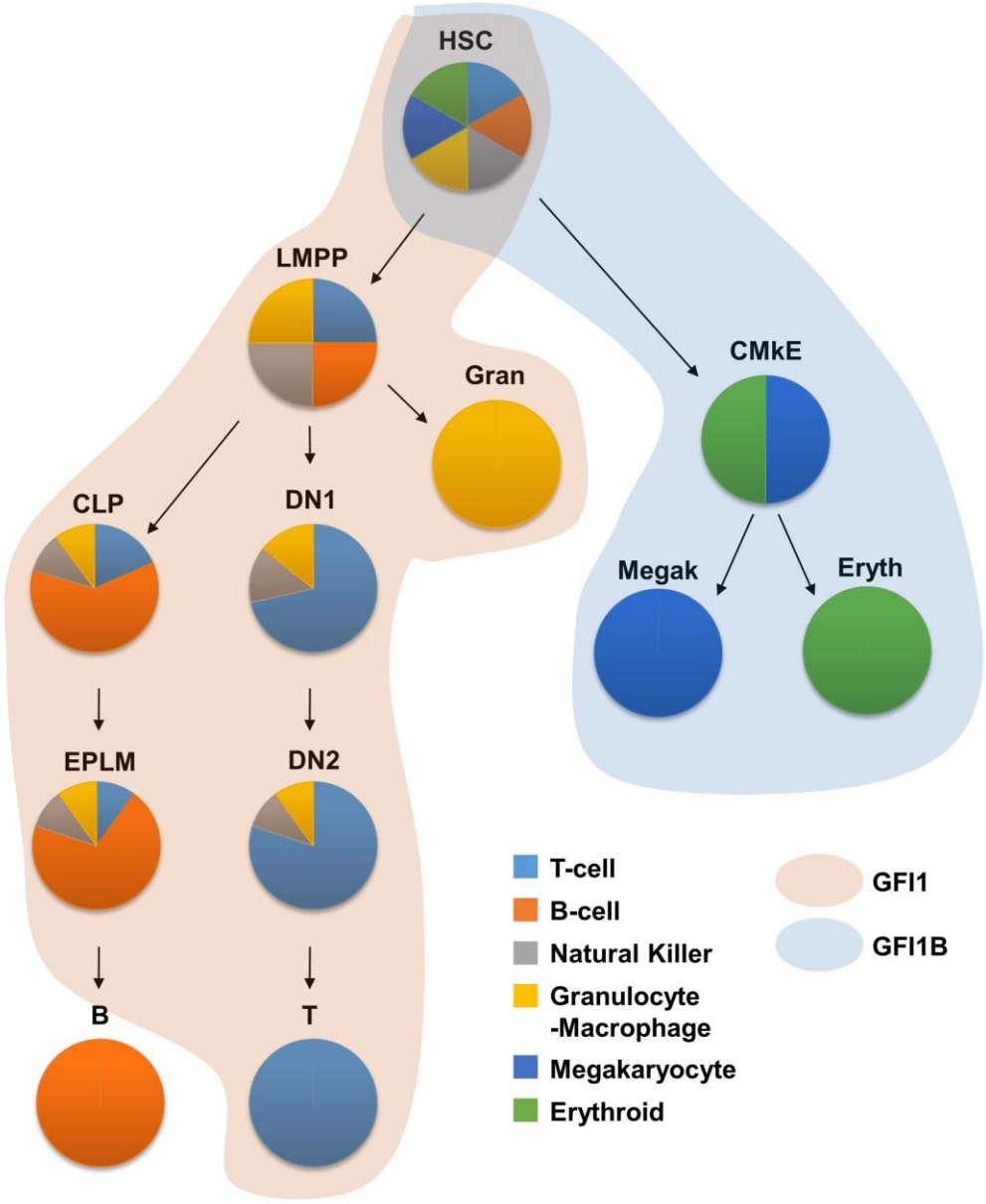

HSC
LMPP
Gran
CMkE
CLP
DN1
Megak
Eryth
EPLM
DN2
B
T
T-cell
B-cell
Natural Killer
Granulocyte
-Macrophage
Megakaryocyte
Erythroid
GFI1
GFI1B

Table 1.1. Abbreviated list of nonhistone proteins regulated by lysine methylation.

Protein:	Motif:	Residue:	Methyltransferase:	Demethylase:
p53	^{370}KSKK^{373}	K370	SMYD2	LSD1
	^{370}KSKK^{373}	K372	SETD7	
	^{370}KSKK^{373}	K373	G9a Glp	
	^{379}RHKK^{382}	K382	SETD8	
RELA/p65	^{36}YKCE^{39}	K37	SETD7	
	^{217}DKVQK^{221}	K218	NSD1	
	^{217}DKVQK^{221}	K221	NSD1	
	^{310}KSIMKK^{315}	K310	Glp	
	^{310}KSIMKK^{315}	K314	SETD7	
	^{310}KSIMKK^{315}	K315	SETD7	
STAT3	^{137}VTEKQ^{141}	K140	SETD7	LSD1
TAF10	^{186}KSKD^{190}	K189	SETD7	
TAF7	^{2}KSKD^{6}	K5	SETD7	
ERα	^{265}LKHKR^{269}	K266	SMYD2	LSD1
	^{265}LKHKR^{269}	K268	SETD7	LSD1
	^{299}KRSKK^{303}	K302	SETD7	
GLI3	^{433}KRSKIK^{438}	K436	SETD7	
	^{592}DRAK^{595}	K595	SETD7	
AR	^{631}KLKK^{634}	K632	SETD7	
PPP1R12A	^{439}GLRK^{442}	K432	SETD7	LSD1
E2F1	^{182}KKSK^{185}	K185	SETD7	LSD1
RB	^{809}LKSPYK^{813}	K810	SMYD2	
	^{858}VLKR^{861}	K860	SMYD2	
	^{870}KPLKK^{874}	K873	SETD7	
RARα	^{107}GCKG^{110}	K109		
	^{167}KEVPK^{171}	K171		
GATA4	^{298}YMKL^{301}	K300	EZH2	
MYOD	^{99}KACKRK^{104}	K104	G9a	
FOXO3	^{268}AKKKA^{272}	K270	SETD7	
	^{268}AKKKA^{272}	K271	SETD7	
KLF12	^{311}SRKR^{314}	K313	G9a	
IRF1	^{124}KSKS^{128}	K126	SETD7	
NR1H4	^{204}KSKR^{207}	K206	SETD7	
MTA1	^{530}VRKP^{533}	K532	G9a	
MEF2D	^{265}SRKP^{268}	K267	G9a	
HSP90	^{207}VKKH^{210}	K209	SMYD2	
PCNA	^{247}LKYY^{250}	K248	SETD8	
PARP1	^{527}LKGG^{530}	K528	SMYD2	

CHAPTER 2

SUMOYLATION REGULATES GROWTH FACTOR INDEPENDENCE 1 IN TRANSCRIPTIONAL CONTROL AND HEMATOPOIESIS

 Andrade D, Velinder M, Singer J, Maese L, Bareyan D, Nguyen H, Chandrasekharan M, Lucente H, McClellan D, Jones D, Sharma S, Liu F, Engel ME (2016) SUMOylation Regulates Growth Factor Independence 1 in Transcriptional Control and Hematopoiesis. Molecular and Cellular Biology 10.1128/MCB.01001-15

AMERICAN SOCIETY FOR MICROBIOLOGY | Molecular and Cellular Biology

SUMOylation Regulates Growth Factor Independence 1 in Transcriptional Control and Hematopoiesis

Daniel Andrade,[a*] Matthew Velinder,[a] Jason Singer,[a] Luke Maese,[b,c] Diana Bareyan,[a] Hong Nguyen,[d] Mahesh B. Chandrasekharan,[e] Helena Lucente,[a] David McClellan,[a] David Jones,[a*] Sunil Sharma,[g,h] Fang Liu,[f] Michael E. Engel[a,b,c,i]

Department of Oncological Sciences, University of Utah School of Medicine, Salt Lake City, Utah, USA[a]; Department of Pediatrics, University of Utah School of Medicine, Salt Lake City, Utah, USA[b]; Primary Children's Hospital, Salt Lake City, Utah, USA[c]; Department of Pediatrics, Vanderbilt University School of Medicine, Nashville, Tennessee, USA[d]; Department of Radiation Oncology, University of Utah School of Medicine, Salt Lake City, Utah, USA[e]; Susan Lehman Cullman Laboratory for Cancer Research, Department of Chemical Biology, Ernest Mario School of Pharmacy, Rutgers Cancer Institute of New Jersey, Rutgers, The State University of New Jersey, Piscataway, New Jersey, USA[f]; Department of Medicine, University of Utah School of Medicine, Salt Lake City, Utah, USA[g]; Center for Investigational Therapeutics, Huntsman Cancer Institute, Salt Lake City, Utah, USA[h]; Nuclear Control Program, Huntsman Cancer Institute, Salt Lake City, Utah, USA[i]

Cell fate specification requires precise coordination of transcription factors and their regulators to achieve fidelity and flexibility in lineage allocation. The transcriptional repressor growth factor independence 1 (GFI1) is comprised of conserved Snail/Slug/Gfi1 (SNAG) and zinc finger motifs separated by a linker region poorly conserved with GFI1B, its closest homolog. Moreover, GFI1 and GFI1B coordinate distinct developmental fates in hematopoiesis, suggesting that their functional differences may derive from structures within their linkers. We show a binding interface between the GFI1 linker and the SP-RING domain of PIAS3, an E3-SUMO (small ubiquitin-related modifier) ligase. The PIAS3 binding region in GFI1 contains a conserved type I SUMOylation consensus element, centered on lysine-239 (K239). *In silico* prediction algorithms identify K239 as the only high-probability site for SUMO modification. We show that GFI1 is modified by SUMO at K239. SUMOylation-resistant derivatives of GFI1 fail to complement Gfi1 depletion phenotypes in zebrafish primitive erythropoiesis and granulocytic differentiation in cultured human cells. LSD1/CoREST recruitment and *MYC* repression by GFI1 are profoundly impaired for SUMOylation-resistant GFI1 derivatives, while enforced expression of MYC blocks granulocytic differentiation. These findings suggest that SUMOylation within the GFI1 linker favors LSD1/CoREST recruitment and *MYC* repression to govern hematopoietic differentiation.

The molecular machinery governing multipotential cell fate is generally characterized by flexibility and combinatorial diversity to achieve distinct outcomes from a limited collection of tools. This principle is observed commonly in hematopoiesis, where factors may contribute to maintenance of stem cells and uncommitted progenitors yet also may be critical determinants of terminal differentiation. The growth factor independence (GFI) family of transcriptional regulators, comprised of *GFI1* and *GFI1B*, typifies this principle.

GFI1 was identified in a Moloney murine leukemia virus (MoMuLV) insertional mutagenesis screen for factors conferring interleukin-2 (IL-2)-independent growth upon cultured lymphocytes (1). *GFI1B* was discovered from its homology to *GFI1* (2, 3). Both proteins harbor nearly identical Snail/Slug/Gfi1 (SNAG) domains and a highly conserved concatemer of six C_2H_2-type zinc fingers (ZnFs) at their N and C termini, respectively. Linker regions with limited conservation separate the SNAG and ZnF motifs (4). SNAG domains confer recruitment of LSD1 and its binding partner CoREST, which figure prominently in transcriptional repression by GFI proteins (5). Sequence-specific DNA binding to a common response element [TA**AATC**AC(A/T)GCA; response element in boldface] is coordinated by ZnFs 3, 4, and 5 in both proteins (2, 6, 7), while ZnFs 1, 2, and 6 and the linkers provide interacting surfaces for transcriptional partners and epigenetic regulators. Despite extraordinary conservation within their SNAG and ZnF regions, GFI1 and GFI1B support distinct fates in hematopoietic differentiation (4, 8, 9).

GFI1 regulates hematopoietic stem cell (HSC) self-renewal and maintains HSC quiescence (10–12). Early stages in B-cell and T-cell lymphopoiesis (B- and T-lymphopoiesis, respectively) and T-cell subset allocation in the adaptive immune response also require *GFI1* (13, 14). Within the myelo-erythroid compartment, *GFI1* is required for both qualitative and quantitative aspects of granulocyte development. Two dominant mutations in GFI1, N382S and K403R, cause severe congenital neutropenia (SCN) type 2 (15, 16), while compound heterozygosity for these mutations has been described in cyclic neutropenia (17). Additionally, reduced GFI1 expression cooperates with mutant C/EBPε in specific granule deficiency (SGD) (18). Outside the hematopoietic compartment, *GFI1* plays crucial roles in sensorineural, neuroendocrine, and intestinal secretory lineage development (19–21). *Gfi1*$^{-/-}$ mice are viable and notably recapitulate these human phenotypes (14, 19, 21–23). In contrast, *Gfi1b* expression appears largely limited to hematopoietic tissues, and its effects are complementary or mutually exclusive to those of *Gfi1* (24). *Gfi1b* is

Received 4 November 2015 **Returned for modification** 8 December 2015
Accepted 20 February 2016

Accepted manuscript posted online 7 March 2016

Citation Andrade D, Velinder M, Singer J, Maese L, Bareyan D, Nguyen H, Chandrasekharan MB, Lucente H, McClellan D, Jones D, Sharma S, Liu F, Engel ME. 2016. SUMOylation regulates growth factor independence 1 in transcriptional control and hematopoiesis. Mol Cell Biol 36:1438–1450. doi:10.1128/MCB.01001-15.

Address correspondence to Michael E. Engel, michael.engel@hci.utah.edu.

* Present address: Daniel Andrade, University of Oklahoma Health Sciences Center, Oklahoma City, Oklahoma, USA; David Jones, Oklahoma Medical Research Foundation, Oklahoma City, Oklahoma, USA.

D.A. and M.V. contributed equally to this article.

required for erythroid and megakaryocyte fate specification, and nullizygosity for *Gfi1b* is lethal by embryonic day 15 (E15) due to failure of definitive erythropoiesis (25). Enforced expression of GFI1B induces T-cell lymphopenia (T-lymphopenia) and impairs granulocytic differentiation *in vitro* (2). These findings suggest that GFI1 and GFI1B can have oppositional roles in lineage allocation (26–28).

Features that functionally distinguish GFI1 from GFI1B are poorly understood, and their distinct roles in hematopoiesis are not completely explained by their patterns of expression. Reporter mouse strains reveal complex expression patterns for both factors (24, 26). During hematopoiesis, both coexpression and lineage-restricted, mutually exclusive expression are observed for *Gfi1* and *Gfi1b*, suggesting that there are points in development when cells must distinguish between them. For example, despite coexpression in HSCs, *Gfi1b* does not complement the stem cell defect in $Gfi1^{-/-}$ mice (10). Likewise, hematopoietic defects in $Gfi1^{-/-}$ mice are only partially reversed by GFI1B expression from the endogenous *Gfi1* promoter, and the defect in sensorineural development is unaffected (23). These findings imply a surprising complexity in systems governing GFI family member function and suggest that elements of primary structure unique to each protein may provide a platform for differential regulation and function.

The human genome encodes three small ubiquitin-related modifier (SUMO) proteins (29, 30) that can be attached to target substrates to modulate their functions. SUMOylation regulates diverse cellular processes, including protein turnover, nucleocytoplasmic transport, subnuclear organization, chromatin structure, and transcriptional control. SUMOs are covalently linked to lysine ε-amino groups in target substrates through coordinated actions of three enzymes (29, 31). An E1 SUMO-activating enzyme links the SUMO C terminus to its active-site cysteine and then transfers SUMO to UBC9, the single E2 SUMO-conjugating enzyme. E3 SUMO ligases then facilitate SUMO transfer to target proteins. Prominent among E3 SUMO ligases are PIAS (protein inhibitor of activated STAT) family proteins (32). PIAS proteins contain a SIZ/PIAS (SP)-RING domain, analogous to the RING domain of E3 ubiquitin ligases, as well as predicted SUMO-interacting motifs (SIM) (29, 33, 34). Through simultaneous binding of UBC9, its bound SUMO, and the target protein, E3 ligases confer substrate and SUMO paralog specificity. Approximately 75% of SUMO-conjugated lysines are found within type I consensus elements, ψKX(D/E), where ψ is a large hydrophobic residue and X is any amino acid (35–37).

A growing roster of transcriptional regulators is being recognized as targets for SUMO conjugation, either via confirmed SUMO attachments or inferred from binding relationships with the SUMOylation machinery (29, 30). We show that GFI1 is modified by SUMO. Mutation of lysine-239 (K239) within the GFI1 linker profoundly limits SUMO modification. GFI1 K239 exists within a type I SUMOylation consensus element embedded in its PIAS3 binding site. Reciprocal mapping studies reveal a binding interface between the PIAS3 SP-RING domain and the C-terminal half of the GFI1 linker. K239-dependent SUMO modification impacts GFI1's role in transcriptional repression and supports both zebrafish primitive erythropoiesis and granulocyte differentiation in HL-60 cells in response to all-*trans* retinoic acid (ATRA). We show that *MYC* expression declines during granulocytic differentiation yet remains elevated with GFI1 depletion. Wild-type GFI1 expression restores both *MYC* suppression and granulocytic differentiation to HL-60 cells depleted of GFI1, yet SUMOylation-resistant GFI1 derivatives fail to do so. Likewise, SUMOylation-resistant GFI1 is impaired in LSD1 and CoREST binding, which figures prominently in GFI1-mediated transcriptional repression. These results define a structural motif in GFI1 that alters its recruitment of LSD1/CoREST to govern its transcriptional repressor function in hematopoietic development.

MATERIALS AND METHODS

Reagents and antibodies. Mouse monoclonal anti-Flag (M2), antitubulin (clone B-5-1-2), anti-PIAS3 (clone PIA3), rabbit anti-Flag (F7425), normal mouse serum (NMS), and normal goat serum (NGS) were obtained from Sigma. Mouse monoclonal anti-myc (9E10) was acquired from Vanderbilt University Medical Center and rabbit polyclonal anti-myc (A-14) was obtained from Santa Cruz Biotechnology, Inc. Mouse monoclonal anti-green fluorescent protein (anti-GFP) (7.1/13.1) was obtained from Roche Applied Science, Inc. Rabbit polyclonal anti-SUMO (ab30258) and anti-GFI1 (ab55108) antibodies were bought from Abcam. Horseradish peroxidase (HRP)-conjugated anti-mouse, anti-rabbit, anti-rat, and anti-goat IgG were obtained from Jackson ImmunoResearch. Alexa Fluor 488-conjugated anti-rabbit IgG was purchased from ThermoFisher Scientific, and anti-CoREST monoclonal antibody (D612U) was purchased from Cell Signaling Technologies. Lipofectamine, Mirus, and To-Pro-3 iodide were obtained from Invitrogen. Protein G-Sepharose was purchased from Sigma-Aldrich. Fugene-6 and a Dual-Luciferase assay kit were purchased from Promega. HCI-2509 synthesis, purification, and characterization have been previously described (38). Puromycin (puro), hygromycin (hygro), *o*-dianisidine, and Polybrene were purchased from Sigma. Restriction endonucleases, polymerases, and ligases were purchased from New England Biosciences. All other materials were of reagent grade.

Plasmids and subcloning. Expression plasmid pCS3-6×-myc:PIAS3α (with six copies of the myc tag) and its derivative fragments consisting of PIAS3 amino acids 1 to 273, 274 to 584, 274 to 392, and 393 to 584 have been previously described (39). Plasmids expressing Flag-tagged PIAS1, PIASxα, PIASxβ, PIAS3α, PIASy, HDAC1, and β-catenin have been described previously (39, 40). The GFI1:3×Flag (with three copies of the Flag tag) expression vector was a generous gift from Tarik Moroy and has been previously described (28, 41). The pcDNA3.1$^+$-6×-myc expression vector was created by subcloning the 6×myc coding sequence, bounded by BamH1 and EcoR1 restriction sites, into BamH1/EcoR1-restricted pcDNA3.1$^+$. Rat *Gfi1* cDNA was subcloned into EcoR1/XbaI-restricted pcDNA3.1$^+$-6×-myc. cDNAs for N-terminal and C-terminal deletion (ΔN and ΔC, respectively) derivatives of rat *Gfi1* were generated by PCR. Amplimers were restricted and subcloned into pcDNA3.1+-6×-myc. The K239R and E241Q substitutions in rat GFI1 were created by two-stage PCR and splicing by overlap extension (42). Expression constructs for 3×Flag-tagged SUMO1, SUMO2, and SUMO3 were generated by PCR and subcloned into pCMV5 and eMIG′+. Retroviruses expressing GFI1:3×Flag, GFI1-K239R:3×-Flag, and GFI1-E241Q:3×Flag were produced using Gateway technology with the pDONR221 and pMIP.RFA plasmids. Retroviruses expressing Flag-tagged SUMO and GFI1 constructs were collected from the supernatants of transfected HEK293T cells. The pEIZ-*MYC* and pEIT vector control were generous gifts of Don Ayer. All constructs were confirmed by automated dideoxy sequencing at the University of Utah sequencing core facility. Primer sequences and further details of subcloning strategies used to generate constructs are available upon request.

Cell culture, transient transfections, and retroviral transduction. COS7L, HEK293T (American Type Culture Collection), and 293-T-REx-5×Gal-luciferase cells (generously provided by Raphael Margueron) were maintained in Dulbecco's modified Eagle's medium (DMEM) supplemented with 10% fetal bovine serum (FBS). NIH 3T3 fibroblasts were propagated in DMEM supplemented with 10% bovine calf serum (BCS). HL-60 cells were maintained in Iscove's modified Dulbecco's medium (IMDM) supplemented with 20% FBS. All cell culture media were sup-

plemented with 2 mM L-glutamine, 50 units/ml penicillin, and 50 μg/ml streptomycin. Transient transfections in COS7L cells were performed with Lipofectamine per the manufacturer's instructions. For retrovirus production in HEK293T cells, monolayers were transfected using Fugene-6 or Mirus, and then supernatants containing viral particles were collected on consecutive days and clarified by centrifugation. HL-60 cells were transduced with retroviral particles by successive spinoculation in the presence of 8 μg/ml Polybrene. Selection in 4 μg/ml puromycin was conducted to generate stable cells. For differential RNA interference (RNAi) experiments, HL-60 cells were transduced with pMKO1-puro expressing a *GFI1*-targeted small hairpin RNA (shRNA) or a scrambled control and selected in puromycin to establish stable isolates. Expression of Flag-tagged GFI1 or its K239R or E241Q derivatives was restored in GFI1-depleted HL-60 cells by retroviral transduction using a pMSCV-hygro (where MSCV is murine stem cell virus) vector containing the appropriate cDNAs, followed by selection in hygromycin.

Immune precipitation and immunoblotting. Following enforced expression of the proteins indicated in the figure legends or those expressed endogenously, immune precipitation and/or immunoblotting was performed essentially as described previously using antibodies indicated in figure legends (43).

Transcriptional reporter assays. 293-T-REx-5×Gal-luciferase cells were transfected with pRL and expression plasmids for Gal4 fusion constructs as shown in figures and referred to in the figure legends. LSD1 inhibitor HCI-2509 was employed at 100 nM in the indicated experiments. Luciferase activities were determined essentially as described previously (44). Statistical significance was determined via an unpaired Student *t* test in GraphPad Prism, version 6.0.

TCA precipitation/SUMOylation assays. COS7L cells expressing Flag-tagged GFI1, GFI1-K239R, GFI1-E241Q, β-catenin, or HDAC1 were harvested in ice-cold phosphate-buffered saline (PBS) and collected by centrifugation. Cell pellets were washed with PBS and then snap-frozen in liquid nitrogen (LN_2). Cell pellets were thawed into 20% trichloroacetic acid (TCA) and subjected to sonication. Precipitating material was collected by centrifugation and then solubilized in 1× SDS-PAGE buffer supplemented with 0.5 mg/ml bovine serum albumin. This solution was neutralized to pH 7.5 with Tris base and then boiled. Flag-tagged proteins were analyzed by SDS-PAGE and immunoblotting or were collected by immune precipitation with rabbit anti-Flag antibody and protein G-Sepharose beads. Immune complexes were fractionated by SDS-PAGE and then subjected to immunoblotting.

Immunofluorescence immunohistochemistry. NIH 3T3 fibroblasts stably expressing Flag-tagged GFI1 or GFI1-K239R were grown on coverslips, fixed in buffered 3.7% formalin, washed with PBS, made permeable with 1% Triton X-100, again washed with PBS, and blocked with 10% NGS in PBS. GFI1 or GFI1-K239R was detected with rabbit anti-Flag followed by Alexa Fluor 488-conjugated goat anti-rabbit secondary antibody and epifluorescence microscopy. Nuclei were revealed with To-Pro-3 iodide.

Morpholino injection. Endogenous *gfi1aa* depletion in zebrafish was achieved using a splice-blocking, antisense morpholino oligonucleotide (5′-CCAATCTAGCCTGAAAATGGCACAA-3′) (Gene Tools, LLC) targeting the intron 1/exon 2 boundary of the primary transcript. Morpholinos were injected at the one-cell stage using 1 nl of a 20 μM solution. Wild-type and mutant rescue constructs were subcloned into pCS2+, and RNA was synthesized through *in vitro* mMessage mMachine transcription reactions (Ambion, Inc.). Rescue RNA constructs were coinjected with the morpholino at 100 ng/μl. Expression of GFI1 and its derivatives in zebrafish embryos was determined by Western blotting directed at the Flag epitope tag at the C terminus of each protein. Experiments with zebrafish were conducted under University of Utah IACUC approved protocol 14-07007.

Reverse transcriptase PCR. Total RNAs were isolated from zebrafish embryos, treated as shown in Fig. 4, at 48 h postfertilization (hpf) using Trizol and reverse transcribed. Primers hybridized within intron 1 (sense) and exon 2 (antisense) to amplify a 169-bp sequence spanning the intron 1-exon 2 boundary in unspliced *gfi1aa* mRNA. A second sense primer in exon 1 paired with the exon 2 antisense primer was used to amplify a 263-bp fragment in spliced *gfi1aa* mRNA. Primer sequences are available by request.

***o*-Dianisidine staining.** Hemoglobinized cells of the developing zebrafish were visualized using *o*-dianisidine staining. Briefly, embryos were collected and dechorionated by pronase digestion at 3 mg/ml. Embryos were anesthetized with 0.2% tricaine for 40 min. Embryos at 48 hpf were stained in the dark on a rolling rotisserie for 2 h in a freshly prepared solution composed of 0.6 mg/ml *o*-dianisidine in 40% ethanol (EtOH), 0.65% H_2O_2, and 0.01 M sodium acetate. Embryos were thoroughly washed with PBS-Tween 20 (PBS-T) and photographed on an Olympus SZX16 dissecting microscope equipped with an Olympus DP71 camera. Scores from 1 to 4, representing none to complete hemoglobinization, respectively, were assigned for 200 embryos by an investigator blinded to the experimental conditions, and statistical significance was determined by Wilcoxon-Mann-Whitney testing in GraphPad Prism, version 6.0.

HL-60 cell differentiation. Naive HL-60 cells, scrambled controls, cells having undergone GFI1 depletion with or without rescue with GFI1 or its K239R or E241Q derivative, or those with enforced MYC expression were treated with 0.1 μM ATRA or vehicle for 4 days. Aliquots were removed during treatment to analyze CD11B cell surface expression by flow cytometry and at the end of the experiment to confirm expression from rescue constructs. On day 4, cells were transferred to glass slides by cytospin, subjected to Wright staining, and scored in a blinded fashion for morphological maturation. Promyelocyte, myelocyte, and metamyelocyte morphologies were scored as immature, while cells whose nuclei displayed band form or segmented morphologies were scored as mature. In parallel, total RNA was collected from stable isolates using an RNeasy minikit (Qiagen). Changes in *CEBPA*, *MYC*, and *AZU* gene expression levels were assessed by quantitative reverse transcription-PCR (qRT-PCR) using the $\Delta\Delta C_T$ (where C_T is threshold cycle) method with β-glucuronidase gene (*GUS*) expression as an internal control and normalized to untreated vector controls. Statistical significance was assessed by an unpaired Student's *t* test using GraphPad Prism, version 6.0. Primer sequences employed for RT-PCR analysis of GFI1 target genes are available upon request.

RESULTS

PIAS3 SP-RING domain binds the GFI1 linker. Through its interaction with PIAS3, GFI1 reverses the inhibitory effect of PIAS3 toward STAT3-dependent transcription (45). To further our understanding of the GFI1-PIAS3 relationship, we first confirmed the GFI1-PIAS3 interaction (Fig. 1A) (45) and then defined the boundaries of their binding interface using transiently expressed, myc-tagged GFI1 or deletion derivatives with Flag-tagged PIAS3 by immune precipitation and Western blotting (Fig. 1B and C). An interaction between full-length GFI1 and PIAS3 was clearly visible. The SNAG and zinc finger domains of GFI1 were dispensable for PIAS3 binding. Instead, a motif between residues 136 and 254 of GFI1 was required to bind PIAS3. Additional contributions were made by the N-terminal half of the linker as deleting residues 21 to 135 diminished PIAS3 binding. However, this region was neither necessary nor sufficient for GFI1 to interact with PIAS3.

To gain additional insights into the relationship between GFI1 and PIAS3, we employed a collection of fragments (Fig. 1B) spanning the PIAS3 primary structure to map the GFI1 binding site (39). Myc-tagged PIAS3 fragments (Fig. 1D) were transiently expressed in COS7L cells with Flag-tagged GFI1, and coprecipitating PIAS3 fragments were detected in anti-Flag immune complexes by anti-myc Western blotting. A C-terminal fragment of PIAS3 containing the SP-RING domain, acidic domain (AD), and the

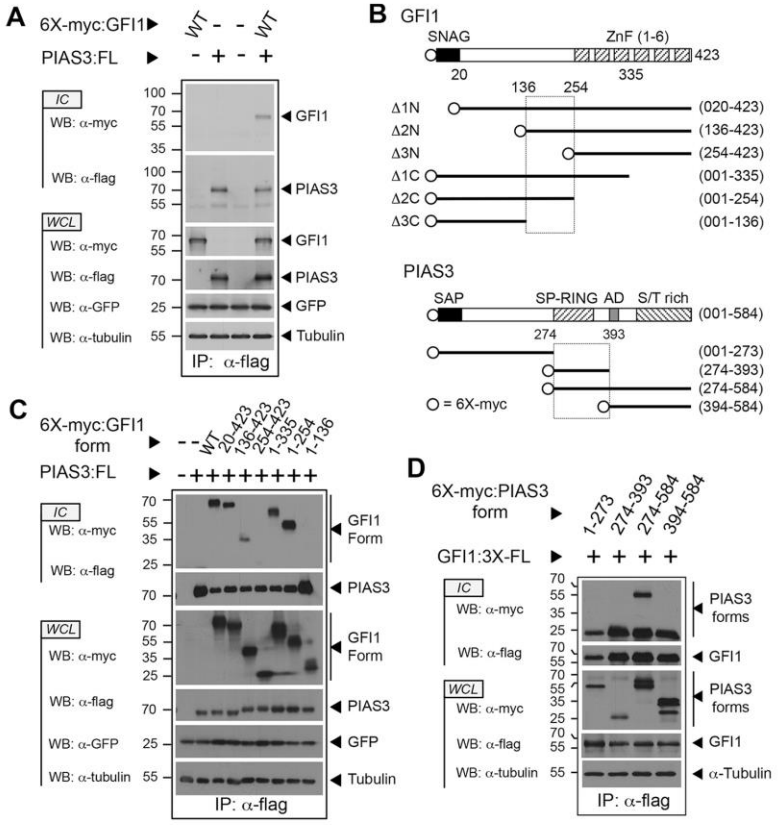


FIG 1 The C-terminal half of the GFI1 linker binds the PIAS3 SP-RING domain. (A) PIAS3 binds GFI1. Myc-tagged GFI1 and Flag-tagged PIAS3 were expressed as shown. PIAS3 was immunopurified from whole-cell lysates (WCL), and coprecipitation of GFI1 was determined by immunoblot (Western blotting [WB]) analysis of immune complexes (IC). Green fluorescent protein (GFP) and tubulin were employed as controls for transfection efficiency and gel loading, respectively. (B) Structures of GFI1 and PIAS3 derivatives employed in the experiments shown in panels C and D. SNAG and ZnF regions are shown for GFI1. SAP, SP-RING, AD, and S/T-rich domains are shown for PIAS3. Amino acid residues corresponding to fragment boundaries are indicated in parentheses. Hatched boxes represent regions forming the GFI1-PIAS3 binding interface. Open circles indicate the 6×myc tag. (C and D) Mapping the GFI1-PIAS3 binding interface. PIAS3 and GFI1 forms, with accompanying epitope tags, were expressed as shown. Coprecipitating GFI1 (C) and PIAS3 (D) forms were identified by Western blotting. WT, wild type; FL, Flag epitope tag; IP, immunoprecipitation.

serine/threonine (S/T)-rich tail readily bound GFI1. Deleting the SP-RING domain (amino acids 274 to 322) and 71 additional residues before the AD completely abolished GFI1 binding. We also observed interactions between GFI1 and PIASxα, PIASxβ, PIASy, and PIAS1 (data not shown). Primary structure alignment for PIAS proteins revealed rapidly declining homology past the SP-RING domain, suggesting that the conserved SP-RING domain of PIAS family proteins contributes significantly to GFI1 binding.

GFI1 SUMOylation involves a SUMO consensus motif within its linker region. Interactions between GFI1 and PIAS family E3 SUMO ligases suggest that GFI1 could be modified by SUMO. To address this hypothesis, we first screened for high-molecular-weight (HMW) forms of GFI1 captured under instantly denaturing conditions, suggesting conjugation by ubiquitin family proteins. COS7L cells transiently expressing Flag-tagged GFI1 were snap-frozen in LN_2, thawed into 20% trichloroacetic acid (TCA), and made soluble by boiling in buffer containing 1% SDS. Solubilized proteins were analyzed by immunoblotting with anti-Flag antibody to detect HMW GFI1 derivatives (Fig. 2A). In parallel, Flag-tagged HDAC1 and β-catenin were expressed as controls for SUMOylation and ubiquitination, respectively. Both controls showed evidence of anti-Flag immune reactivity in the HMW range, suggesting that both SUMO- and ubiquitin-modified targets could be visualized by this technique. For GFI1, a prominent HMW population was observed, as well as several GFI1 fragments consistent with proteolysis. To determine if HMW forms of GFI1 contained SUMO and/or ubiquitin, we purified Flag-tagged GFI1 from proteins solubilized after isolation in TCA as described above. GFI1-containing immune complexes subjected to Western blotting with anti-SUMO and antiubiquitin antibodies clearly showed SUMO and ubiquitin conjugation (Fig. 2B).

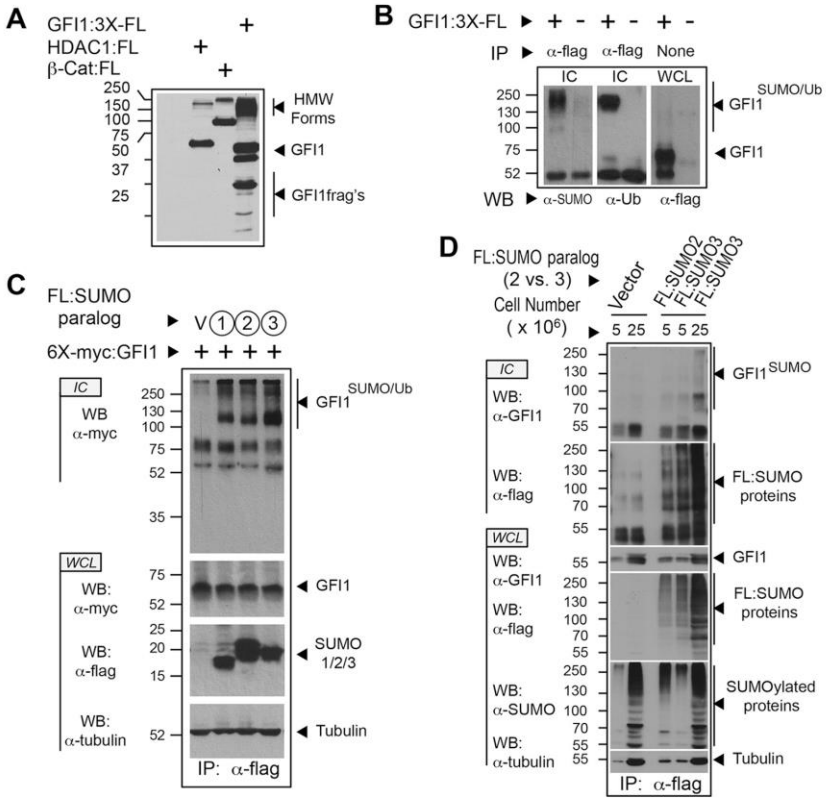


FIG 2 GFI1 is SUMOylated and ubiquitinated. (A) High-molecular-weight (HMW) derivatives of GFI1 identified under instantly denaturing conditions. COS7L cells were transfected with empty vector or expression constructs for Flag-tagged GFI1, HDAC1, and β-catenin. Cells were collected in-ice cold PBS, snap-frozen in liquid nitrogen, and suspended in 20% TCA. Precipitated proteins were solubilized by boiling in solubilizing buffer. Aliquots were analyzed by immunoblotting with anti-Flag antibody M2. (B) HMW derivatives of GFI1 contain SUMO and ubiquitin modifications. Flag-tagged GFI1 was expressed in COS7L cells, collected, and processed as described for panel A. Anti-Flag immune complexes (IC) and whole-cell lysate (WCL) were analyzed by SDS-PAGE and Western blotting (WB) as shown. (C) Each SUMO paralog can conjugate GFI1. Flag-tagged SUMO1, SUMO2, or SUMO3 was expressed in COS7L cells with myc-tagged GFI1 as shown. Cells were harvested as described for panel A. SUMOylated proteins were collected by anti-Flag immune precipitation. Immune complexes and whole-cell lysates were analyzed by Western blotting (WB). (D) Endogenously expressed GFI1 is SUMOylated. HL-60 cells in the quantities shown, stably expressing Flag-tagged SUMO2, SUMO3, or vector control, were collected and processed as described for panel A and then immunopurified via the Flag epitope tag (FL). Immune complexes (IC) were probed for GFI1 by Western blotting (WB). Expression of GFI1 and SUMO proteins was confirmed in whole-cell lysates (WCL) by Western blotting. Ub, ubiquitin.

The apparent molecular masses of HMW GFI1 derivatives approach 250 kDa, suggesting that GFI1 may be conjugated by multiple SUMO and/or ubiquitin moieties. Among SUMO family proteins, only SUMO2 and SUMO3 can form poly-SUMO chains, owing to the presence of a SUMOylation consensus element in both proteins (30). SUMO1 lacks a SUMOylation consensus motif, enabling mono-SUMOylation and restricting its presence in poly-SUMO chains to the terminal SUMO modification (30). To address the possibility that GFI1 can be modified by SUMO1, -2, or -3, we expressed Flag-tagged versions of each SUMO with myc-tagged GFI1, isolated SUMOylated targets by anti-Flag immune precipitation from TCA-precipitated, resolubilized proteins, and subjected immune complexes to anti-myc Western blotting. GFI1 conjugation by each SUMO paralog could be clearly demonstrated (Fig. 2C). To confirm SUMOylation at endogenous levels of GFI1 expression, we created HL-60 cells stably expressing Flag-tagged SUMO2 or SUMO3 and immunopurified SUMO2- or SUMO3-conjugated proteins and probed for GFI1 by Western blotting (Fig. 2D). As expected, GFI1 signal was detected in the HMW range for both SUMO2- and SUMO3-conjugated proteins. Similar results were observed for HL-60 cells stably expressing SUMO1 (data not shown). These findings indicate that endogenously expressed GFI1 can be modified by each SUMO paralog and collectively suggest an intimate relationship between GFI1 and the SUMOylation machinery.

GFI1 SUMO modification requires K239 within an intact SUMOylation consensus element. Approximately 75% of experimentally confirmed SUMO modifications occur within ψKX

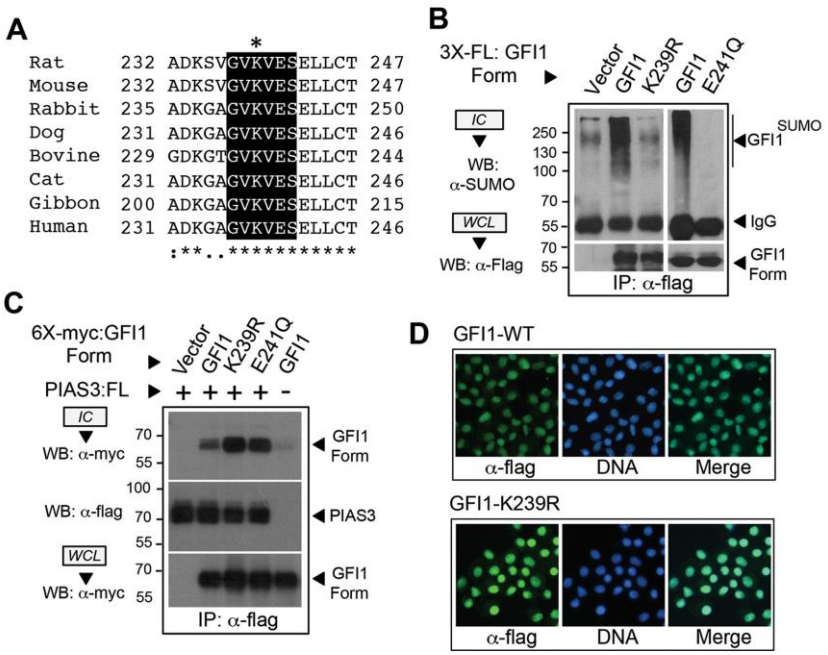


FIG 3 GFI1 SUMOylation requires a type I SUMOylation consensus element in the GFI1 linker. (A) Alignment and conservation of a type I SUMOylation consensus sequence in the GFI1 linker in selected mammalian species. The putative SUMO acceptor lysine (K239 in rat) is indicated by an asterisk. (B) SUMO modification of GFI1 is abolished by K239R or E241Q substitution. Flag-tagged GFI1 or its K239R or E241Q derivative was expressed and then isolated by anti-Flag immune precipitation (IP) from total cellular protein harvested under instantly denaturing conditions as described in the legend to Fig. 2. Immune complexes (IC) and whole-cell lysates (WCL) were subjected to immunoblotting with anti-SUMO or anti-Flag antibodies as shown. (C) GFI1 and its K239R or E241Q derivative display PIAS3 binding. A 6×myc-tagged GFI1 or its K239R or E241Q derivative was expressed with Flag-tagged PIAS3. Presence of GFI1, GFI1-K239R, or GFI1-E241Q in anti-Flag immune complexes was determined by Western blotting. (D) GFI1 and its K239R derivative localize to the nucleus. NIH 3T3 cells were transduced with retrovirus expressing Flag-tagged GFI1 or its K239R derivative, and then subcellular localization was determined in stable polyclonal populations by epifluorescence detection. Nuclei were counterstained with To-Pro-3-iodide.

(D/E) SUMOylation consensus elements. The GFI1 linker contains a motif (GVKVES; motif is underlined) within the PIAS3 binding site, centered on K239, that matches a type I SUMOylation consensus sequence (Fig. 3A). This motif is absolutely conserved in mammalian GFI1 orthologs (46). Yet given imperfect prediction by the SUMOylation consensus alone, we utilized three *in silico* prediction algorithms of SUMOylation, SUMOplot (Abgent) (47), SUMOsp2.0 (37), and SUMOhydro (36), to identify high-probability sites for SUMO conjugation in GFI1 (data not shown). Among GFI1 lysine residues, only K239 met the criteria of being conserved among mammalian GFI1 orthologs, being absent from GFI1B, and being identified as a high-probability site in each algorithm. To assess K239 in SUMO modification of GFI1, we created an arginine substitution at this site (K239R). Because type I SUMOylation consensus elements have an invariant aspartate or glutamate at the +2 position, we also created a glutamate (E)-to-glutamine (Q) substitution at E241 (E241Q), leaving the acceptor lysine (K239) intact. Using instantly denaturing LN_2/TCA/SDS conditions followed by anti-Flag immune precipitation and anti-SUMO immunoblotting, SUMO modification of GFI1 was readily observed. However, neither the K239R nor E241Q derivative of GFI1 showed appreciable SUMO conjugation (Fig. 3B). Notably, GFI1 and its K239R and E241Q derivatives displayed comparable PIAS3 binding and nuclear localizations (Fig. 3C and D), implying that SUMO conjugation failure was not due to altered subcellular localization or interaction with PIAS3. These data indicate that GFI1 SUMOylation depends upon K239 within an intact, type I SUMOylation consensus element.

GFI1 SUMOylation supports cell fate determination in hematopoiesis. GFI1 serves multiple roles in developmental hematopoiesis, involving both myeloid and lymphoid lineages. The *GFI1* homolog in zebrafish, *gfi1aa*, is required for primitive erythropoiesis (48). We leveraged this requirement to determine the impact of GFI1 SUMOylation *in vivo*. Hemoglobinized cells characteristic of primitive erythropoiesis were readily revealed by *o*-dianisidine staining of zebrafish embryos at 48 hpf (Fig. 4A). To render observations quantitative across a continuum, we used a scoring system to represent none to complete hemoglobinization, with scores assigned to each of 200 embryos by an investigator blinded to the experimental conditions. Using a morpholino oligonucleotide designed to block splicing of the *gfi1aa* mRNA (Fig. 4B), we confirmed that *gfi1aa* depletion could abort primitive erythropoiesis and that wild-type rat *Gfi1* could complement this deficiency. However, despite expression comparable to that of

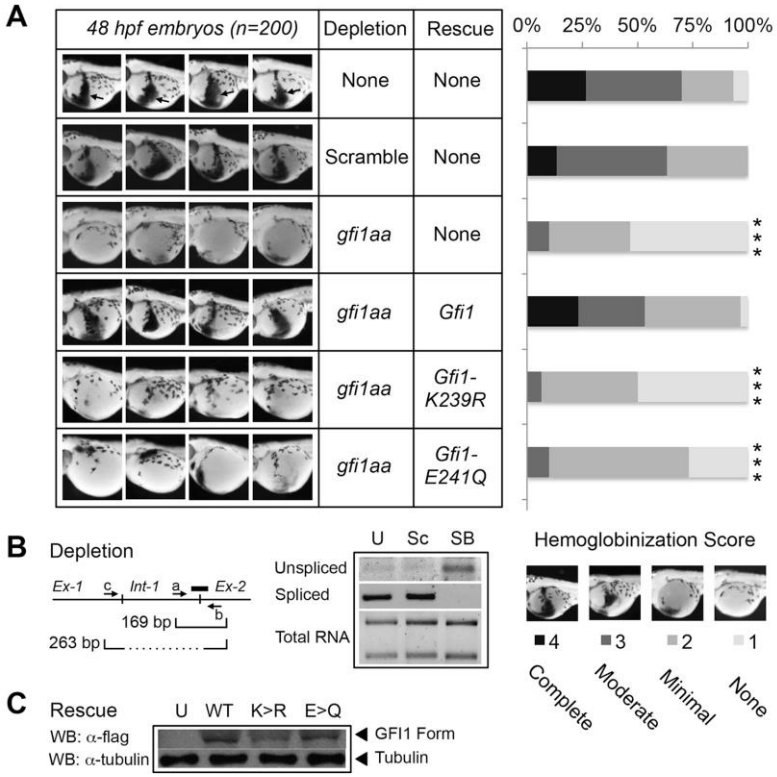


FIG 4 GFI1 SUMOylation on K239 is required for zebrafish primitive erythropoiesis. (A) One-cell-stage zebrafish embryos were injected with *gfi1aa* splice-blocking morpholino, alone or in combination with RNA expressing wild-type GFI1, GFI1-K239R, or GFI1-E241Q, each with a C-terminal Flag epitope tag. At 48 h postfertilization (hpf), primitive erythropoiesis was revealed by *o*-dianisidine staining. Four representative embryos are shown out of 200 embryos scored for each knockdown/rescue combination. Arrows in uninjected controls show the zone of primitive erythropoiesis. Primitive erythropoiesis was quantified using a graded scoring system. Scores of 1 to 4 were assigned for each embryo to indicate none, modest, moderate, or complete hemoglobinization, respectively. Statistical significance was assessed using a Wilcoxon-Mann-Whitney test (***, $P < 0.0005$). (B) Morpholino-induced splicing blockade of *gfi1aa* in zebrafish. Total RNA was isolated from uninjected (U) zebrafish embryos or after injection of a scrambled control (Sc) or *gfi1aa* splice-blocking morpholino (SB) targeting the boundary between intron 1 and exon 2, as shown (thick line). Unspliced *gfi1aa* cDNA was amplified with primers a and b spanning the intron 1-exon 2 boundary to yield a 169-bp amplimer. Spliced *gfi1aa* mRNA was amplified using primers c and b to yield a 263-bp amplimer. (C) Expression of Flag-tagged GFI1, GFI1-K239R (K>R), and GFI1-E241Q (E>Q) derivatives in extracts prepared from the equivalent of eight zebrafish embryos at 24 h postinjection.

wild-type GFI1 (Fig. 4C) neither Gfi1-K239R nor Gfi1-E241Q could do so (29, 30). These data indicate that GFI1 SUMOylation supports primitive erythropoiesis in zebrafish.

GFI1 is also required for mammalian neutrophil differentiation *in vivo*, and mutations in *GFI1* cause SCN type 2. To extend our findings to mammalian hematopoiesis, we leveraged granulocytic differentiation of HL-60 human promyelocytic leukemia cells in response to all-*trans* retinoic acid (ATRA). HL-60 cells were transduced with retroviruses expressing either a *GFI1*-targeted shRNA or a scrambled control, and GFI1 depletion in stable isolates was confirmed by Western blotting (Fig. 5A). Stable cells were exposed to either ATRA or vehicle and then examined for granulocytic differentiation by CD11B expression (Fig. 5B) and cell morphology (Fig. 5C and D). CD11B expression rose spontaneously in transduced HL-60 cells with time in culture, increased significantly in response to ATRA in control cells, and was reduced in the context of GFI1 depletion. However, when assessed morphologically, granulocytic differentiation was nearly absent with GFI1 depletion. To determine the contribution of GFI1 SUMOylation to this phenotype, we rescued expression with GFI1 or GFI1-K239R in *GFI1*-depleted HL-60 cells, treated stable isolates with ATRA for 4 days, and then assessed granulocytic differentiation morphologically (Fig. 5E). Restoring GFI1 expression complemented the defect in ATRA-mediated granulocyte differentiation, but GFI1-K239R failed to do so despite comparable levels of expression. Collectively, these findings point toward an essential role for GFI1 SUMOylation in hematopoietic differentiation.

GFI1 SUMOylation regulates *MYC* expression during granulocytic differentiation. GFI1 is transactivated by CAAT/enhancer binding protein α (C/EBPα), a critical determinant of granulocytic differentiation, and both GFI1 and C/EBPα are im-

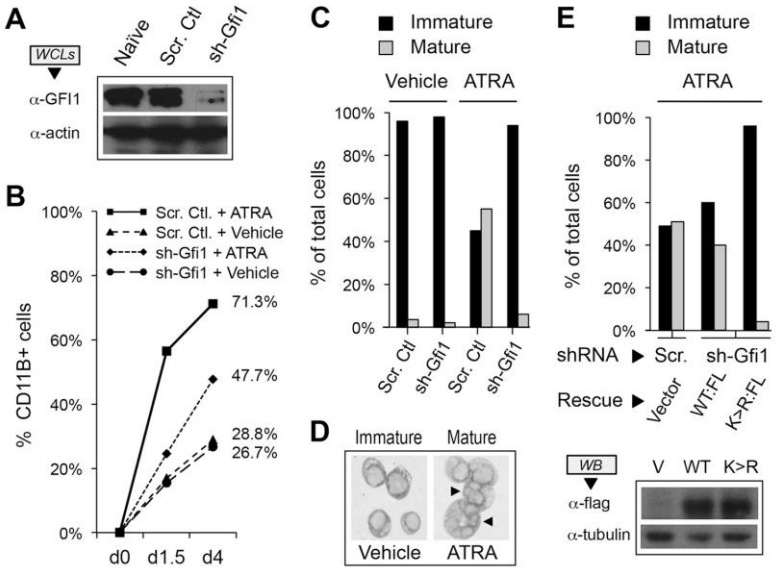


FIG 5 GFI1 SUMOylation supports HL-60 cell granulocytic differentiation in response to all-*trans* retinoic acid (ATRA). (A) GFI1 depletion from HL-60 cells. Whole-cell extracts (WCLs) were prepared from either naive HL-60 cells or those infected with retrovirus expressing small hairpin RNA targeting *GFI1* (sh-Gfi1) or a content-matched scrambled control shRNA (Scr. Ctl.). GFI1 levels were determined by Western blotting. Actin levels were used to confirm equal loading. (B) CD11B expression following stimulation with ATRA. HL-60 cells were treated with 0.1 μM ATRA or vehicle following shRNA-mediated depletion of *GFI1* versus a content-matched scrambled control shRNA. $CD11B^+$ cells were determined by flow cytometry. (C) Granulocytic differentiation of HL-60 cells requires GFI1. HL-60 cells were treated with ATRA or vehicle for 4 days following shRNA-mediated depletion of *GFI1* or a scrambled control. One thousand cells from randomly selected fields for each condition were visually scored as immature (promyelocyte, myelocyte, or metamyelocyte morphology) or mature (band form or multisegmented nuclei). (D) Morphology of immature versus mature HL-60 cells following ATRA treatment. Arrows indicate mature cells with segmented nuclei reminiscent of granulocyte differentiation. (E) GFI1-K239R expression fails to complement the granulocyte maturation defect brought on by *GFI1* depletion. HL-60 cells were subjected to scrambled control or GFI1-depleting shRNA, then rescued with expression constructs for Flag-tagged wild type (WT:FL) or GFI1-K239R (K>R:FL) as shown. After 4 days of ATRA exposure, cells were scored visually for granulocyte maturation as described for panel C. Expression of Flag-tagged GFI1 variants was confirmed by Western blotting.

plicated in repression of *MYC* (49–51). Moreover, *MYC* expression declines precipitously during ATRA-mediated differentiation of HL-60 cells, and enforced expression of MYC blocks terminal differentiation in multiple settings (52–54). To gain mechanistic insights for GFI1 SUMOylation in granulocytic differentiation, we assessed expression of *CEBPA*, *MYC*, and azurocidin gene (*AZU*) in HL-60 cells with and without GFI1, in the absence or presence of ATRA (Fig. 6A). Expression changes for these factors in HL-60 cells treated with ATRA and occupancy of their promoters by GFI1 have been previously described (52). HL-60 cells were transduced with retroviruses expressing either a *GFI1*-targeted shRNA or a scrambled control, and then GFI1 expression was restored relative to the level with empty vector rescue. *CEBPA* expression did not change appreciably with GFI1 depletion or enforced GFI1 expression and fell modestly under each condition following ATRA exposure. *MYC* expression fell precipitously with ATRA exposure in control cells but rose significantly with GFI1 depletion, and in response to ATRA *MYC* expression did not fall below the level seen for undifferentiated controls. Moreover, *MYC* repression was restored by enforced expression of GFI1 and further potentiated by ATRA exposure. *AZU* expression fell in HL-60 cells exposed to ATRA, failed to do so in GFI1-depleted cells, but was again repressed when GFI1 expression was restored. To address contributions from GFI1 SUMOylation to transcriptional control, we again depleted HL-60 cells of GFI1, restored expression of GFI1 or its SUMOylation-deficient forms, K239R and E241Q, and then assessed expression of *CEBPA*, *MYC*, and *AZU* in the absence or presence of ATRA (Fig. 6B). None of these GFI1 forms showed a significant impact upon *CEBPA* expression. However, while GFI1 restored repression of *MYC* and *AZU* in response to ATRA, the K239R and E241Q derivatives of GFI1 were impaired. These data indicate that GFI1-mediated repression of *MYC* and *AZU* correlates with its effects on granulocytic differentiation and that SUMOylation favors this function of GFI1.

In multiple settings, enforced expression of MYC impairs terminal differentiation (53, 54). To determine the impact of elevated *MYC* expression resulting from GFI1 depletion, we expressed MYC constitutively in HL-60 cells and then assessed granulocytic differentiation in response to ATRA relative to that of naive and vector control cells (Fig. 7). While granulocytic maturation was readily apparent in naive and vector control cells following ATRA exposure, it was noticeably impaired in the context of enforced MYC expression. These data suggest that increases in MYC expression can block HL-60 cell differentiation in response to ATRA and support the possibility that *MYC* repression by

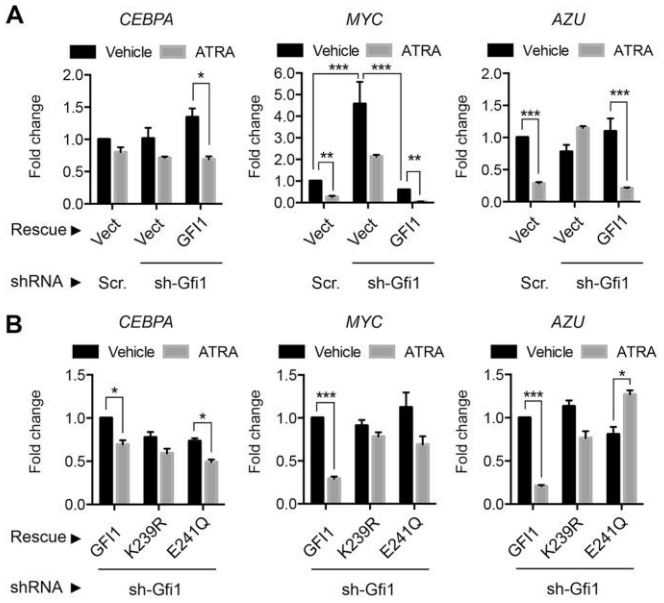


FIG 6 SUMOylation modulates GFI1-dependent *MYC* expression to direct ATRA-mediated granulocyte maturation in HL-60 cells. (A) GFI1 modulates *MYC* expression in ATRA-mediated granulocytic differentiation. *GFI1*-targeted shRNA (sh-Gfi1) or a content-matched scrambled control (Scr) was used to deplete *GFI1* in HL-60 cells, followed by rescue with a vector control (Vect) or GFI1 as shown. Levels of *CEBPA*, *MYC*, and *AZU* (azurocidin) mRNAs were determined by qRT-PCR relative to that of the *GUS* internal control. Fold change in expression is shown relative to that of the untreated, scrambled or vector control cells. (B) GFI1 was depleted from HL-60 cells using *GFI1*-targeted shRNA, followed by restored expression of GFI1, GFI1-K239R, or GFI1-E241Q and then treatment with vehicle or ATRA as shown. Expression of *CEBPA*, *MYC*, and *AZU* was measured as described for panel A. Statistical significance was determined by Wilcoxon-Mann-Whitney testing (*, $P < 0.05$; **, $P < 0.005$; ***, $P < 0.0005$).

SUMOylation-competent GFI1 enables this differentiated phenotype.

SUMOylation favors GFI1-LSD1/CoREST binding and transcriptional repression. LSD1 binds the GFI1 SNAG domain, and factors that disrupt GFI1-LSD1 binding impair GFI1-mediated transcriptional repression (4, 5, 55). Similarly, the demethylase activity of LSD1 toward mono- and dimethylated histone H3 lysine 4 (H3K4me1/2) in nucleosomes is augmented by its interaction with CoREST and is further favored by SUMO2 and SUMO3 (SUMO2/3) binding to a SIM in CoREST (56–58). These findings suggest that the SUMOylation of transcription factors that partner with LSD1-CoREST may favor LSD1 activity at target promoters. To explore contributions from SUMOylation to transcriptional repression by GFI1, we used the reporter cell line 293-T-REx-5×Gal-TK-luciferase (59–61) and fusion proteins comprised of GFI1 sequences fused to the DNA binding region of Gal4 (Fig. 8A). Eliminating GFI1 ZnFs 4 and 5 prevents binding at GFI1-regulated endogenous promoters, instead directing the GFI1 repressor function to the 5×Gal-upstream activation sequence (UAS) of the integrated reporter construct. Thus, the GFI1 repressor function on endogenous promoters can be assessed in a chromatinized context independent of partnerships with other DNA binding proteins. In this model system, GFI1-Δ1C:Gal4 displayed dose-dependent transcriptional repression. SNAG-WT:Gal4 (where WT is wild type) and GFI1-Δ1C/K239R:Gal4 also showed dose-dependent reductions in reporter output that were comparable but moderately and consistently less than that of GFI1-Δ1C:Gal4. This difference became less apparent at higher plasmid doses. These findings mirror those showing significant but submaximal transcriptional repression attributable to the

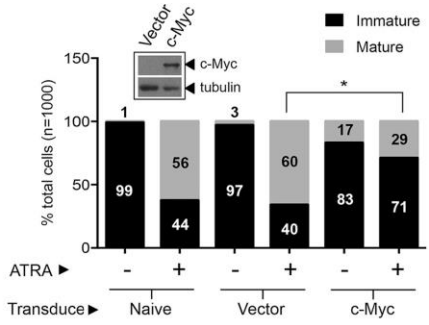


FIG 7 Enforced MYC expression blocks granulocytic differentiation of HL-60 cells in response to ATRA. Naive HL-60 cells and those transduced with a MYC-expressing retrovirus or with a control vector were treated with 0.1 μM ATRA for 4 days and scored for immature versus mature granulocytic morphology. The Western blot shows c-Myc expression in transduced cells relative to that in vector control cells (inset) using anti-myc rabbit polyclonal antibody, A-14. Tubulin serves as a loading control. Statistical significance was determined by Wilcoxon-Mann-Whitney testing (*, $P < 0.05$).

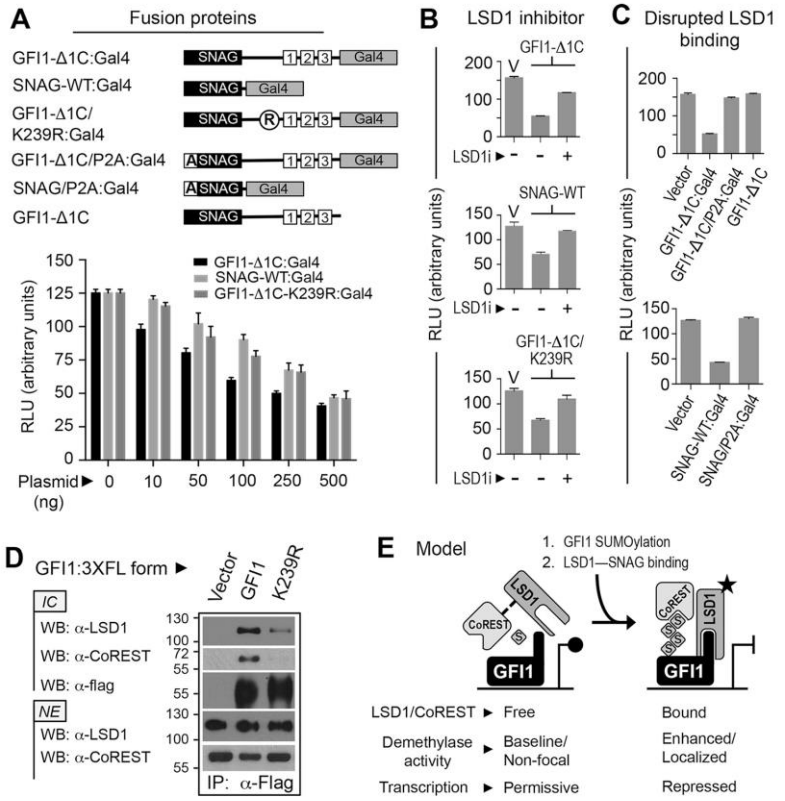


FIG 8 SUMOylation supports LSD1/CoREST binding and transcriptional repression by GFI1. (A) SUMOylation contributes to GFI1-mediated repression. GFI1-Δ1C, the SNAG domain, or GFI1-Δ1C with the K239R substitution was expressed as a fusion with Gal4 in HEK293-T-REx-5×Gal-luciferase cells grown in six-well plates. Firefly luciferase activity was determined relative to that of *Renilla* luciferase using a dual-luciferase assay kit. (B and C) LSD1 activity contributes to transcriptional repression by GFI1. GFI1-Δ1C, the SNAG domain, or GFI1-Δ1C/K239R was expressed as a Gal4 fusion protein in the presence or absence of LSD1 inhibitor, HCI-2509 (LSD1i). Luciferase activity was scored as described for panel A. Similarly, transcriptional repression by GFI1-Δ1C or the SNAG domain, each harboring the P2A mutation that abolishes LSD1 binding, was compared to repression produced by its wild-type configurations at this position. (D) GFI1-K239R displays impaired LSD1/CoREST binding. Flag-tagged GFI1 or GFI1-K239R was expressed in COS7L cells and immunopurified from nuclear extracts (NE) by anti-Flag immune precipitation. Endogenously expressed LSD1 and CoREST coprecipitating with GFI1 or GFI1-K239R were determined by Western blotting (WB) of anti-Flag immune complexes (IC). (E) A model hypothesizing how GFI1 SUMOylation could influence transcriptional repression by the GFI1-LSD1/CoREST complex. RLU, relative light units.

GFI1 SNAG domain in prior studies (55). Repression was antagonized by the LSD1 inhibitor HCI-2509 (Fig. 8B) and by the P2A substitution that abolishes SNAG-LSD1 binding (Fig. 8C) (38). These data indicate that LSD1 recruitment dominates transcriptional repression by GFI1, that SUMOylation supports this repressor function, and that when SUMOylation is impaired, repression by GFI1 resembles that achieved by the SNAG domain alone.

The GFI1 SNAG domain interacts directly with LSD1 and indirectly recruits its corepressor, CoREST (5). In light of impaired transcriptional repression by GFI1-K239R, we hypothesized impaired LSD1/CoREST binding by this mutant. In coprecipitation assays, we found that endogenously expressed LSD1 and CoREST interacted readily with wild-type GFI1 (Fig. 8D). However, despite its intact SNAG domain, GFI1-K239R was noticeably impaired in LSD1 and CoREST binding. These findings indicate that GFI1 SUMOylation favors LSD1/CoREST binding, providing a functional link between SUMO conjugation and transcriptional repression.

DISCUSSION

GFI1 is a transcriptional repressor that controls growth, differentiation, and survival in normal and malignant hematopoiesis. Yet we have only a limited understanding of factors governing GFI1 function. By focusing on the GFI1 linker region, we have identified a type I SUMOylation motif unique to GFI1 and embedded within its PIAS3 binding site. K239R and E241Q substitutions within this site abolish GFI1 SUMOylation, and unlike wild-type GFI1, these derivatives fail to complement GFI1 depletion phenotypes. Likewise, these SUMOylation-deficient derivatives are impaired in

transcriptional repression at GFI1 target genes. These data suggest that SUMOylation at K239 favors GFI1-mediated transcriptional control and cell fate determination.

It is likely that SUMOylation occurs directly on K239 as the E241Q substitution is similarly impaired yet retains the K239 ε-amino group for alternative posttranslational modifications. However, these functional defects could also reflect modification(s) elsewhere in GFI1 that depends upon K239 and/or E241, including conjugation by SUMO or other ubiquitin family members. SUMOylation on multiple lysines has been indicated for other transcriptional regulators, including c-Jun, p53, and Nkx2.5 (62, 63). Notably, studies of a 25-amino-acid peptide from SENS, the *Drosophila melanogaster* ortholog of GFI1, show that K509 can be SUMOylated *in vitro*. Yet a K509R derivative of SENS is still SUMOylated when expressed in S2 cells (64), indicating alternative sites for SUMO conjugation. K509 is analogous to K403 in GFI1, which is mutated in SCN type 2 (15, 16), suggesting that a defect in K403 SUMOylation might impair GFI1 function in neutrophil development. However, K403 in GFI1 does not occupy a consensus SUMOylation site, nor has K403 SUMOylation ever been demonstrated. Defining the full complement of GFI1 modifications by SUMO/ubiquitin family members and the contexts in which these modifications occur will be important to understanding their contributions to GFI1-mediated transcriptional control and cell fate specification.

Our findings indicate that GFI1 SUMOylation correlates with transcriptional repression, as has been seen for other DNA binding proteins (65–67). LSD1 is the dominant effector of GFI1-mediated transcriptional repression. Its demethylase activity toward H3K4me1/2 peptide substrates is enhanced by CoREST, while CoREST-LSD1 binding favors demethylation of H3K4me1/2 in nucleosomes. Moreover, an interaction between SUMO2/3 and the SIM in CoREST potentiates its activation of LSD1 demethylase activity. SUMO1 does not do so. Since LSD1 function at target promoters should be proportional to its recruitment and its activation state, factors that favor either should support transcriptional repression while those that compromise either should impair it. We observed a prominent decline in both LSD1 and CoREST binding to GFI1-K239R relative to that of GFI1, and this correlated with impaired transcriptional repression and complementation failure in functional assays. Based upon these findings, we speculate that SUMO-CoREST binding provides an additional tether for LSD1 bound to the GFI1 SNAG domain and augments LSD1 demethylase activity. Therefore, SUMOylation-deficient derivatives of GFI1 may display relatively less LSD1 binding and activity, causing functional impairment.

The finding that GFI1 can be modified by SUMO1, -2, or -3 and ubiquitin suggests considerable complexity in GFI1 regulation by ubiquitin family members. For example, whether GFI1 is modified by SUMO1 or SUMO2/3 could influence CoREST recruitment and stimulation of intrinsic LSD1 activity at GFI1-regulated promoters. Distinct patterns of SUMO/ubiquitin conjugation could impact composition, activity, and integrity of GFI1 complexes, thereby expanding its functional repertoire. Moreover, SUMO/ubiquitin modifications potentially occur on the same residues under different conditions. These modifications would be mutually exclusive and may also compete with lysine modifications by other small molecules derived from metabolic pathways to modulate gene expression based upon cellular needs (68–70). As such, our results may reflect one of many potential complexes assembled to regulate GFI1-responsive genes, each favored by a discrete constellation of posttranslational modifications.

Changes in gene expression are both energetically costly and impactful. Thus, requiring dual inputs before executing changes seems not only reasonable but advantageous. With respect to the GFI1-LSD1/CoREST axis, LSD1 binding to the SNAG domain would serve as the first input, with GFI1 SUMOylation and CoREST recruitment serving as the second (Fig. 8E). This model predicts that transcriptional repression by GFI1 would be most favored by concurrent SNAG-LSD1 and SUMO2/3-CoREST binding. The former would deliver LSD1 to GFI1-regulated genes while the latter would stabilize LSD1 binding and augment its demethylase activity. Should CoREST engage LSD1 in solution, its stimulating effect toward LSD1 would be limited by the need for a secondary input from SUMO2/3 binding. LSD1 actions could then be limited by reversal of either event or by GFI1 turnover. This fail-safe strategy could help ensure fidelity when a change in gene expression is initiated and yet provide mechanisms to modulate and ultimately terminate it. Thus, SUMO/ubiquitin modifications to GFI1 may offer additional inputs governing transcriptional control and supporting distinct, context-dependent cell fate decisions in hematopoiesis. We are only at the beginning of unraveling the potential regulatory complexity for this deceptively simple transcriptional repressor.

ACKNOWLEDGMENTS

We are indebted to colleagues Srividya Bhaskara, Trudy Oliver, Thomas O'Hare, Don Ayer, Stephen Lessnick, Kimble Frazer, and Nicholas Trede for critical review and helpful suggestions regarding the manuscript. We gratefully acknowledge support from the oligonucleotide synthesis, DNA sequencing, and Centralized Zebrafish Animal Resource (CZAR) core facilities of the University of Utah and the Huntsman Cancer Institute.

The Huntsman Cancer Foundation and the National Institutes of Health under award numbers P30CA042014, K08DK080190 (M.E.E.), and T32DK007115-39 (L.M.) provided support for the research presented in this publication. Additional support was provided via a Career Development Award from the St. Baldrick's Foundation (M.E.E.) and from the Office of the Vice President for Research of the University of Utah School of Medicine.

FUNDING INFORMATION

This work, including the efforts of Michael Eugene Engel, was funded by HHS | NIH | National Cancer Institute (NCI) (P30CA042014-26). This work, including the efforts of Michael Eugene Engel, was funded by HHS | NIH | National Institute of Diabetes and Digestive and Kidney Diseases (NIDDK) (K08DK080190). This work, including the efforts of Luke Maese, was funded by HHS | NIH | National Institute of Diabetes and Digestive and Kidney Diseases (NIDDK) (T32DK007115-39). This work, including the efforts of Michael Eugene Engel, was funded by St. Baldrick's Foundation (Career Development Award).

The funders had no role in the design of experiments, collection of data, or interpretation of study results. Likewise, they had no role in the decision to submit the work for publication

REFERENCES

1. **Gilks CB, Bear SE, Grimes HL, Tsichlis PN.** 1993. Progression of interleukin-2 (IL-2)-dependent rat T cell lymphoma lines to IL-2-independent growth following activation of a gene (Gfi-1) encoding a novel zinc finger protein. Mol Cell Biol **13**:1759–1768. http://dx.doi.org/10.1128/MCB.13.3.1759.
2. **Tong B, Grimes HL, Yang TY, Bear SE, Qin Z, Du K, El-Deiry WS, Tsichlis PN.** 1998. The Gfi-1B proto-oncoprotein represses $p21^{WAF1}$ and

inhibits myeloid cell differentiation. Mol Cell Biol **18:**2462–2473. http://dx.doi.org/10.1128/MCB.18.5.2462.

3. **Rodel B, Wagner T, Zornig M, Niessing J, Moroy T.** 1998. The human homologue (GFI1B) of the chicken GFI gene maps to chromosome 9q34.13-A locus frequently altered in hematopoietic diseases. Genomics **54:**580–582. http://dx.doi.org/10.1006/geno.1998.5601.
4. **van der Meer LT, Jansen JH, van der Reijden BA.** 2010. Gfi1 and Gfi1b: key regulators of hematopoiesis. Leukemia **24:**1834–1843. http://dx.doi.org/10.1038/leu.2010.195.
5. **Saleque S, Kim J, Rooke HM, Orkin SH.** 2007. Epigenetic regulation of hematopoietic differentiation by Gfi-1 and Gfi-1b is mediated by the cofactors CoREST and LSD1. Mol Cell **27:**562–572. http://dx.doi.org/10.1016/j.molcel.2007.06.039.
6. **Zweidler-Mckay PA, Grimes HL, Flubacher MM, Tsichlis PN.** 1996. Gfi-1 encodes a nuclear zinc finger protein that binds DNA and functions as a transcriptional repressor. Mol Cell Biol **16:**4024–4034. http://dx.doi.org/10.1128/MCB.16.8.4024.
7. **Lee S, Doddapaneni K, Hogue A, McGhee L, Meyers S, Wu Z.** 2010. Solution structure of Gfi-1 zinc domain bound to consensus DNA. J Mol Biol **397:**1055–1066. http://dx.doi.org/10.1016/j.jmb.2010.02.006.
8. **Phelan JD, Shroyer NF, Cook T, Gebelein B, Grimes HL.** 2010. Gfi1-cells and circuits: unraveling transcriptional networks of development and disease. Curr Opin Hematol **17:**300–307. http://dx.doi.org/10.1097/MOH.0b013e32833a06f8.
9. **Kazanjian A, Gross EA, Grimes HL.** 2006. The growth factor independence-1 transcription factor: new functions and new insights. Crit Rev Oncol Hematol **59:**85–97. http://dx.doi.org/10.1016/j.critrevonc.2006.02.002.
10. **Zeng H, Yucel R, Kosan C, Klein-Hitpass L, Moroy T.** 2004. Transcription factor Gfi1 regulates self-renewal and engraftment of hematopoietic stem cells. EMBO J **23:**4116–4125. http://dx.doi.org/10.1038/sj.emboj.7600419.
11. **Hock H, Hamblen MJ, Rooke HM, Schindler JW, Saleque S, Fujiwara Y, Orkin SH.** 2004. Gfi-1 restricts proliferation and preserves functional integrity of haematopoietic stem cells. Nature **431:**1002–1007. http://dx.doi.org/10.1038/nature02994.
12. **Liu Y, Elf SE, Miyata Y, Sashida G, Liu Y, Huang G, Di Giandomenico S, Lee JM, Deblasio A, Menendez S, Antipin J, Reva B, Koff A, Nimer SD.** 2009. p53 regulates hematopoietic stem cell quiescence. Cell Stem Cell **4:**37–48. http://dx.doi.org/10.1016/j.stem.2008.11.006.
13. **Spooner CJ, Cheng JX, Pujadas E, Laslo P, Singh H.** 2009. A recurrent network involving the transcription factors PU.1 and Gfi1 orchestrates innate and adaptive immune cell fates. Immunity **31:**576–586. http://dx.doi.org/10.1016/j.immuni.2009.07.011.
14. **Yucel R, Karsunky H, Klein-Hitpass L, Moroy T.** 2003. The transcriptional repressor Gfi1 affects development of early, uncommitted c-Kit+ T cell progenitors and CD4/CD8 lineage decision in the thymus. J Exp Med **197:**831–844. http://dx.doi.org/10.1084/jem.20021417.
15. **Person RE, Li FQ, Duan Z, Benson KF, Wechsler J, Papadaki HA, Eliopoulos G, Kaufman C, Bertolone SJ, Nakamoto B, Papayannopoulou T, Grimes HL, Horwitz M.** 2003. Mutations in proto-oncogene GFI1 cause human neutropenia and target ELA2. Nat Genet **34:**308–312. http://dx.doi.org/10.1038/ng1170.
16. **Xia J, Bolyard AA, Rodger E, Stein S, Aprikyan AA, Dale DC, Link DC.** 2009. Prevalence of mutations in ELANE, GFI1, HAX1, SBDS, WAS and G6PC3 in patients with severe congenital neutropenia. Br J Haematol **147:**535–542. http://dx.doi.org/10.1111/j.1365-2141.2009.07888.x.
17. **Armistead PM, Wieder E, Akande O, Alatrash G, Quintanilla K, Liang S, Molldrem J.** 2010. Cyclic neutropenia associated with T cell immunity to granulocyte proteases and a double de novo mutation in GFI1, a transcriptional regulator of ELANE. Br J Haematol **150:**716–719. http://dx.doi.org/10.1111/j.1365-2141.2010.08274.x.
18. **Khanna-Gupta A, Sun H, Zibello T, Lee HM, Dahl R, Boxer LA, Berliner N.** 2007. Growth factor independence-1 (Gfi-1) plays a role in mediating specific granule deficiency (SGD) in a patient lacking a gene-inactivating mutation in the C/EBPε gene. Blood **109:**4181–4190. http://dx.doi.org/10.1182/blood-2005-05-022004.
19. **Wallis D, Hamblen M, Zhou Y, Venken KJ, Schumacher A, Grimes HL, Zoghbi HY, Orkin SH, Bellen HJ.** 2003. The zinc finger transcription factor Gfi1, implicated in lymphomagenesis, is required for inner ear hair cell differentiation and survival. Development **130:**221–232. http://dx.doi.org/10.1242/dev.00190.
20. **Kazanjian A, Wallis D, Au N, Nigam R, Venken KJ, Cagle PT, Dickey BF, Bellen HJ, Gilks CB, Grimes HL.** 2004. Growth factor independence-1 is expressed in primary human neuroendocrine lung carcinomas and mediates the differentiation of murine pulmonary neuroendocrine cells. Cancer Res **64:**6874–6882. http://dx.doi.org/10.1158/0008-5472.CAN-04-0633.
21. **Shroyer NF, Wallis D, Venken KJ, Bellen HJ, Zoghbi HY.** 2005. Gfi1 functions downstream of Math1 to control intestinal secretory cell subtype allocation and differentiation. Genes Dev **19:**2412–2417. http://dx.doi.org/10.1101/gad.1353905.
22. **Karsunky H, Zeng H, Schmidt T, Zevnik B, Kluge R, Schmid KW, Duhrsen U, Moroy T.** 2002. Inflammatory reactions and severe neutropenia in mice lacking the transcriptional repressor Gfi1. Nat Genet **30:**295–300. http://dx.doi.org/10.1038/ng831.
23. **Fiolka K, Hertzano R, Vassen L, Zeng H, Hermesh O, Avraham KB, Duhrsen U, Moroy T.** 2006. Gfi1 and Gfi1b act equivalently in haematopoiesis, but have distinct, non-overlapping functions in inner ear development. EMBO Rep **7:**326–333. http://dx.doi.org/10.1038/sj.embor.7400618.
24. **Vassen L, Okayama T, Moroy T.** 2007. Gfi1b:green fluorescent protein knock-in mice reveal a dynamic expression pattern of Gfi1b during hematopoiesis that is largely complementary to Gfi1. Blood **109:**2356–2364. http://dx.doi.org/10.1182/blood-2006-06-030031.
25. **Saleque S, Cameron S, Orkin SH.** 2002. The zinc-finger proto-oncogene Gfi-1b is essential for development of the erythroid and megakaryocytic lineages. Genes Dev **16:**301–306. http://dx.doi.org/10.1101/gad.959102.
26. **Yucel R, Kosan C, Heyd F, Moroy T.** 2004. Gfi1:green fluorescent protein knock-in mutant reveals differential expression and autoregulation of the growth factor independence 1 (Gfi1) gene during lymphocyte development. J Biol Chem **279:**40906–40917. http://dx.doi.org/10.1074/jbc.M400808200.
27. **Doan LL, Porter SD, Duan Z, Flubacher MM, Montoya D, Tsichlis PN, Horwitz M, Gilks CB, Grimes HL.** 2004. Targeted transcriptional repression of Gfi1 by GFI1 and GFI1B in lymphoid cells. Nucleic Acids Res **32:**2508–2519. http://dx.doi.org/10.1093/nar/gkh570.
28. **Vassen L, Fiolka K, Mahlmann S, Moroy T.** 2005. Direct transcriptional repression of the genes encoding the zinc-finger proteins Gfi1b and Gfi1 by Gfi1b. Nucleic Acids Res **33:**987–998. http://dx.doi.org/10.1093/nar/gki243.
29. **Wilkinson KA, Henley JM.** 2010. Mechanisms, regulation and consequences of protein SUMOylation. Biochem J **428:**133–145. http://dx.doi.org/10.1042/BJ20100158.
30. **Gareau JR, Lima CD.** 2010. The SUMO pathway: emerging mechanisms that shape specificity, conjugation and recognition. Nat Rev Mol Cell Biol **11:**861–871. http://dx.doi.org/10.1038/nrm3011.
31. **Hochstrasser M.** 2009. Origin and function of ubiquitin-like proteins. Nature **458:**422–429. http://dx.doi.org/10.1038/nature07958.
32. **Shuai K, Liu B.** 2005. Regulation of gene-activation pathways by PIAS proteins in the immune system. Nat Rev Immunol **5:**593–605. http://dx.doi.org/10.1038/nri1667.
33. **Hecker CM, Rabiller M, Haglund K, Bayer P, Dikic I.** 2006. Specification of SUMO1- and SUMO2-interacting motifs. J Biol Chem **281:**16117–16127. http://dx.doi.org/10.1074/jbc.M512757200.
34. **Stehmeier P, Muller S.** 2009. Phospho-regulated SUMO interaction modules connect the SUMO system to CK2 signaling. Mol Cell **33:**400–409. http://dx.doi.org/10.1016/j.molcel.2009.01.013.
35. **Xu J, He Y, Qiang B, Yuan J, Peng X, Pan XM.** 2008. A novel method for high accuracy sumoylation site prediction from protein sequences. BMC Bioinformatics **9:**8. http://dx.doi.org/10.1186/1471-2105-9-8.
36. **Chen YZ, Chen Z, Gong YA, Ying G.** 2012. SUMOhydro: a novel method for the prediction of sumoylation sites based on hydrophobic properties. PLoS One **7:**e39195. http://dx.doi.org/10.1371/journal.pone.0039195.
37. **Ren J, Gao X, Jin C, Zhu M, Wang X, Shaw A, Wen L, Yao X, Xue Y.** 2009. Systematic study of protein sumoylation: development of a site-specific predictor of SUMOsp 2.0. Proteomics **9:**3409–3412. http://dx.doi.org/10.1002/pmic.200800646.
38. **Sorna V, Theisen ER, Stephens B, Warner SL, Bearss DJ, Vankayalapati H, Sharma S.** 2013. High-throughput virtual screening identifies novel N′-(1-phenylethylidene)-benzohydrazides as potent, specific, and reversible LSD1 inhibitors. J Med Chem **56:**9496–9508. http://dx.doi.org/10.1021/jm400870h.
39. **Long J, Wang G, Matsuura I, He D, Liu F.** 2004. Activation of Smad transcriptional activity by protein inhibitor of activated STAT3 (PIAS3).

Proc Natl Acad Sci U S A **101:**99–104. http://dx.doi.org/10.1073/pnas.0307598100.

40. **Moore AC, Amann JM, Williams CS, Tahinci E, Farmer TE, Martinez JA, Yang G, Luce KS, Lee E, Hiebert SW.** 2008. Myeloid translocation gene family members associate with T-cell factors (TCFs) and influence TCF-dependent transcription. Mol Cell Biol **28:**977–987. http://dx.doi.org/10.1128/MCB.01242-07.
41. **Vassen L, Fiolka K, Moroy T.** 2006. Gfi1b alters histone methylation at target gene promoters and sites of gamma-satellite containing heterochromatin. EMBO J **25:**2409–2419. http://dx.doi.org/10.1038/sj.emboj.7601124.
42. **Horton RM, Cai ZL, Ho SN, Pease LR.** 1990. Gene splicing by overlap extension: tailor-made genes using the polymerase chain reaction. Biotechniques **8:**528–535.
43. **Engel ME, Nguyen HN, Mariotti J, Hunt A, Hiebert SW.** 2010. Myeloid translocation gene 16 (MTG16) interacts with Notch transcription complex components to integrate Notch signaling in hematopoietic cell fate specification. Mol Cell Biol **30:**1852–1863. http://dx.doi.org/10.1128/MCB.01342-09.
44. **Barrett CW, Smith JJ, Lu LC, Markham N, Stengel KR, Short SP, Zhang B, Hunt AA, Fingleton BM, Carnahan RH, Engel ME, Chen X, Beauchamp RD, Wilson KT, Hiebert SW, Reynolds AB, Williams CS.** 2012. Kaiso directs the transcriptional corepressor MTG16 to the Kaiso binding site in target promoters. PLoS One **7:**e51205. http://dx.doi.org/10.1371/journal.pone.0051205.
45. **Rodel B, Tavassoli K, Karsunky H, Schmidt T, Bachmann M, Schaper F, Heinrich P, Shuai K, Elsasser HP, Moroy T.** 2000. The zinc finger protein Gfi-1 can enhance STAT3 signaling by interacting with the STAT3 inhibitor PIAS3. EMBO J **19:**5845–5855. http://dx.doi.org/10.1093/emboj/19.21.5845.
46. **Goujon M, McWilliam H, Li W, Valentin F, Squizzato S, Paern J, Lopez R.** 2010. A new bioinformatics analysis tools framework at EMBL-EBI. Nucleic Acids Res **38:**W695–699. http://dx.doi.org/10.1093/nar/gkq313.
47. **Pichler A.** 2008. Analysis of sumoylation. Methods Mol Biol **446:**131–138. http://dx.doi.org/10.1007/978-1-60327-084-7_9.
48. **Cooney JD, Hildick-Smith GJ, Shafizadeh E, McBride PF, Carroll KJ, Anderson H, Shaw GC, Tamplin OJ, Branco DS, Dalton AJ, Shah DI, Wong C, Gallagher PG, Zon LI, North TE, Paw BH.** 2013. Teleost growth factor independence (*gfi*) genes differentially regulate successive waves of hematopoiesis. Dev Biol **373:**431–441. http://dx.doi.org/10.1016/j.ydbio.2012.08.015.
49. **Lidonnici MR, Audia A, Soliera AR, Prisco M, Ferrari-Amorotti G, Waldron T, Donato N, Zhang Y, Martinez RV, Holyoake TL, Calabretta B.** 2010. Expression of the transcriptional repressor Gfi-1 is regulated by C/EBPα and is involved in its proliferation and colony formation-inhibitory effects in p210BCR/ABL-expressing cells. Cancer Res **70:**7949–7959. http://dx.doi.org/10.1158/0008-5472.CAN-10-1667.
50. **Johansen LM, Iwama A, Lodie TA, Sasaki K, Felsher DW, Golub TR, Tenen DG.** 2001. c-Myc is a critical target for C/EBPα in granulopoiesis. Mol Cell Biol **21:**3789–3806. http://dx.doi.org/10.1128/MCB.21.11.3789-3806.2001.
51. **Duan Z, Zarebski A, Montoya-Durango D, Grimes HL, Horwitz M.** 2005. Gfi1 coordinates epigenetic repression of $p21^{Cip/WAF1}$ by recruitment of histone lysine methyltransferase G9a and histone deacetylase 1. Mol Cell Biol **25:**10338–10351. http://dx.doi.org/10.1128/MCB.25.23.10338-10351.2005.
52. **Duan Z, Horwitz M.** 2003. Targets of the transcriptional repressor oncoprotein Gfi-1. Proc Natl Acad Sci U S A **100:**5932–5937. http://dx.doi.org/10.1073/pnas.1031694100.
53. **Leon J, Ferrandiz N, Acosta JC, Delgado MD.** 2009. Inhibition of cell differentiation: a critical mechanism for MYC-mediated carcinogenesis? Cell Cycle **8:**1148–1157. http://dx.doi.org/10.4161/cc.8.8.8126.
54. **Delgado MD, Leon J.** 2010. Myc roles in hematopoiesis and leukemia. Genes Cancer **1:**605–616. http://dx.doi.org/10.1177/1947601910377495.
55. **Grimes HL, Chan TO, Zweidler-McKay PA, Tong B, Tsichlis PN.** 1996. The Gfi-1 proto-oncoprotein contains a novel transcriptional repressor domain, SNAG, and inhibits G_1 arrest induced by interleukin-2 withdrawal. Mol Cell Biol **16:**6263–6272. http://dx.doi.org/10.1128/MCB.16.11.6263.
56. **Lee MG, Wynder C, Cooch N, Shiekhattar R.** 2005. An essential role for CoREST in nucleosomal histone 3 lysine 4 demethylation. Nature **437:**432–435.
57. **Shi YJ, Matson C, Lan F, Iwase S, Baba T, Shi Y.** 2005. Regulation of LSD1 histone demethylase activity by its associated factors. Mol Cell **19:**857–864. http://dx.doi.org/10.1016/j.molcel.2005.08.027.
58. **Ouyang J, Shi Y, Valin A, Xuan Y, Gill G.** 2009. Direct binding of CoREST1 to SUMO-2/3 contributes to gene-specific repression by the LSD1/CoREST1/HDAC complex. Mol Cell **34:**145–154. http://dx.doi.org/10.1016/j.molcel.2009.03.013.
59. **Sarma K, Margueron R, Ivanov A, Pirrotta V, Reinberg D.** 2008. Ezh2 requires PHF1 to efficiently catalyze H3 lysine 27 trimethylation in vivo. Mol Cell Biol **28:**2718–2731. http://dx.doi.org/10.1128/MCB.02017-07.
60. **Li G, Margueron R, Ku M, Chambon P, Bernstein BE, Reinberg D.** 2010. Jarid2 and PRC2, partners in regulating gene expression. Genes Dev **24:**368–380. http://dx.doi.org/10.1101/gad.1886410.
61. **Mozzetta C, Pontis J, Fritsch L, Robin P, Portoso M, Proux C, Margueron R, Ait-Si-Ali S.** 2014. The histone H3 lysine 9 methyltransferases G9a and GLP regulate polycomb repressive complex 2-mediated gene silencing. Mol Cell **53:**277–289. http://dx.doi.org/10.1016/j.molcel.2013.12.005.
62. **Costa MW, Lee S, Furtado MB, Xin L, Sparrow DB, Martinez CG, Dunwoodie SL, Kurtenbach E, Mohun T, Rosenthal N, Harvey RP.** 2011. Complex SUMO-1 regulation of cardiac transcription factor Nkx2-5. PLoS One **6:**e24812. http://dx.doi.org/10.1371/journal.pone.0024812.
63. **Schmidt D, Muller S.** 2002. Members of the PIAS family act as SUMO ligases for c-Jun and p53 and repress p53 activity. Proc Natl Acad Sci U S A **99:**2872–2877. http://dx.doi.org/10.1073/pnas.052559499.
64. **Powell LM, Chen A, Huang YC, Wang PY, Kemp SE, Jarman AP.** 2012. The SUMO pathway promotes basic helix-loop-helix proneural factor activity via a direct effect on the Zn finger protein senseless. Mol Cell Biol **32:**2849–2860. http://dx.doi.org/10.1128/MCB.06595-11.
65. **Gill G.** 2004. SUMO and ubiquitin in the nucleus: different functions, similar mechanisms? Genes Dev **18:**2046–2059. http://dx.doi.org/10.1101/gad.1214604.
66. **Hay RT.** 2005. SUMO: a history of modification. Mol Cell **18:**1–12. http://dx.doi.org/10.1016/j.molcel.2005.03.012.
67. **Geiss-Friedlander R, Melchior F.** 2007. Concepts in sumoylation: a decade on. Nat Rev Mol Cell Biol **8:**947–956. http://dx.doi.org/10.1038/nrm2293.
68. **Walsh CT, Garneau-Tsodikova S, Gatto GJ, Jr.** 2005. Protein posttranslational modifications: the chemistry of proteome diversifications. Angew Chem Int Ed Engl **44:**7342–7372. http://dx.doi.org/10.1002/anie.200501023.
69. **Zhang Z, Tan M, Xie Z, Dai L, Chen Y, Zhao Y.** 2011. Identification of lysine succinylation as a new post-translational modification. Nat Chem Biol **7:**58–63. http://dx.doi.org/10.1038/nchembio.495.
70. **Chen Y, Sprung R, Tang Y, Ball H, Sangras B, Kim SC, Falck JR, Peng J, Gu W, Zhao Y.** 2007. Lysine propionylation and butyrylation are novel post-translational modifications in histones. Mol Cell Proteomics **6:**812–819. http://dx.doi.org/10.1074/mcp.M700021-MCP200.

CHAPTER 3

GFI1 FUNCTIONS IN TRANSCRIPTIONAL CONTROL AND CELL FATE DETERMINATION REQUIRE SNAG DOMAIN METHYLATION TO RECRUIT LSD1

Abstract

Proper hematopoietic cell fate decisions require coordinated functions of transcription factors, their associated coregulators, and histone modifying enzymes. Growth factor independence 1 (GFI1) is a zinc finger transcriptional repressor and master regulator of normal and malignant hematopoiesis. While several GFI1 interacting proteins have been described, how GFI1 leverages these relationships to carry out transcriptional repression remains unclear. Here we describe a functional axis involving GFI1, SMYD2, and LSD1 that is a critical contributor to GFI1 mediated transcriptional repression. SMYD2 methylates lysine-8 (K8) within a -^{8}KSKK^{11}- motif embedded in the GFI1 SNAG domain. Methylation defective GFI1 SNAG domain lacks repressor function due to failure of LSD1 recruitment and persistence of promoter H3K4 dimethyl marks. Methylation defective GFI1 also fails to complement GFI1 depletion phenotypes in developing zebrafish and lacks progrowth and survival functions in lymphoid leukemia cells. Our data show a discrete methylation event in the GFI1 SNAG domain that facilitates recruitment of LSD1 to enable transcriptional repression and coordinate control of hematopoietic cell fate in both normal and malignant settings.

Introduction

Growth factor independence 1 (GFI1) is a transcriptional repressor and master regulator of normal and malignant hematopoiesis. GFI1 is comprised of a 20 amino acid N-terminal Snail/GFI1 (SNAG) domain, a C-terminal concatemer of six C_2H_2-type zinc fingers and a linker region that separates them (1). GFI1

mediated transcriptional repression requires DNA binding via zinc fingers 3, 4, and 5 (2). The SNAG domain is the dominant contributor to transcriptional repression by GFI1 (2). GFI1 and its paralog, GFI1B, share 95% identity within their SNAG domains and have a high degree of conservation within their zinc fingers (3). This conservation is reflected by their recognition of a common response element [TA**AATC**AC(A/T)GCA; response element in bold face] and shared binding partners, yet knock in experiments in mice show GFI1 and GFI1B only partially substitute for one another in hematopoietic development (4).

GFI1 contributes to hematopoietic stem cell (HSC) quiescence and self renewal (5, 6). GFI1 is also an obligate transcription factor in granulopoiesis and is equally critical to establish and maintain lymphoid fate in development of the adaptive immune system (3, 7). Mutations (N382S, K403R) within GFI1 zinc fingers act in a dominant negative fashion to cause severe congenital neutropenia (SCN) (8, 9). In the malignant setting, GFI1 is required to establish and maintain lymphoid leukemia/lymphoma, at least, in part, by restricting proapoptotic functions of p53 (10-12). These reports highlight the critical roles filled by GFI1 in both normal and malignant hematopoiesis.

GFI1 functions through discrete interactions with partner DNA binding proteins, transcriptional coregulators, and histone modifying enzymes, including ETS1, MTG8, MTG16, G9a, HDAC1, and LSD1 (13-16). Factors influencing GFI1 interactions with these partners to govern transcriptional repression are incompletely understood. By virtue of its demethylase activity toward mono and/or dimethylated histone H3, lysine 4 (H3K4me1/me2), LSD1 is the dominant effector

for GFI1 mediated transcriptional repression and is the only protein known to bind the SNAG domain in either GFI1 or GFI1B (1). A single amino acid substitution from proline to alanine at residue 2 (P2A) within the SNAG domain profoundly impairs LSD1 binding and GFI1 repressor function (2). Likewise, disruption of LSD1 binding to the GFI1B SNAG domain impairs GFI1B mediated gene repression and disrupts proper erythroid and megakaryocytic differentiation (14). Notably, substitutions at nearby residues (R3A, S4A, K10A, and K11A) have minimal impact on GFI1 repressor function (2). These data suggest residue specific impact upon LSD1 binding and GFI1 function via the SNAG domain, yet leaves as an open question how this impact is established and whether GFI1-LSD1 binding and repressor functions might be regulated via modifications at these residues.

The SNAG domain in GFI1 contains a family and species conserved lysine-serine-lysine-lysine (KSKK) motif occurring at residues 8 through 11 (14). Over 300 UniProt genes encode a KSKK motif, including one occurring from residues 370 to 373 in p53. Residue specific lysine methylation in the p53 -^{370}KSKK^{373}- motif, catalyzed by protein lysine methyltransferases SMYD2, SETD7, G9a and Glp, dynamically regulate p53 function (17-20). From this, we hypothesized the GFI1 -^{8}KSKK^{11}- motif may be similarly regulated by methylation, and given functional connections between GFI1 and p53, that the same methyltransferases may govern GFI1 transcriptional repression function attributable to SNAG—LSD1 binding. We show that SMYD2 binds and methylates K8 within the GFI1 SNAG domain to promote its interaction with LSD1. Leucine substitution at K8 (K8L)

abolishes LSD1 binding required for transcriptional repression by the SNAG domain in isolation. Likewise, GFI1-K8L fails to complement the defect in zebrafish primitive erythropoiesis brought on by *gfi1aa* depletion and is unable to support T-ALL cell survival after depletion of endogenously expressed GFI1. These data indicate that SNAG domain methylation governs GFI1—LSD1 axis functions and when considered with other GFI1 posttranslational modifications, may be one of several regulatory inputs to this critical determinant of hematopoietic cell fate decisions.

Materials and methods

Reagents and antibodies

Mouse monoclonal α-FLAG (M2) and α-tubulin (B-5-1-2) antibodies were obtained from Sigma Aldrich. Rabbit polyclonal α-myc (A-14) and goat polyclonal α-GFI1 (sc-8558, N-20) antibodies were obtained from Santa Cruz Biotechnology. Rabbit polyclonal α-H3 (ab1791) and rabbit monoclonal α-H3K4me2 (ab32356) antibodies were obtained from Abcam. Rabbit monoclonal α-LSD1 (C69G12, #2184) antibody was obtained from Cell Signaling Technology. Horseradish peroxidase (HRP) conjugated α-mouse, α-rabbit, and α-goat IgG were obtained from Jackson Immunoresearch. LSD1 inhibitor, HCI-2509, has been described previously (21). Restriction endonucleases, polymerases, and ligases were purchased from New England Biosciences. All other materials were of reagent grade.

Plasmids and subcloning

GFI1:3X-FLAG constructs were a generous gift from Dr. Tarik Moroy (22). P2A and K8L substitutions were generated by two-stage PCR and splicing by overlap extension. GFI1 SNAG domain derivative fusion proteins with the Gal4 DNA binding domain were generated by first annealing complementary single stranded DNA oligonucleotides encoding SNAG domain derivative sense and antisense strands. The sense strand incorporated an EcoRI site at its 5' end while the antisense strand contained a MluI site at its 5' end. Annealed oligonucleotides encoding SNAG domain derivatives were restricted with EcoRI/MluI and subcloned into EcoRI/MluI restricted pCMV5. The Gal4 DNA binding domain sequence (encoding amino acid residues 1-147) was inserted in-frame and 3' of the respective SNAG domain derivative using XbaI/BamHI restriction sites. LSD1 constructs have been previously described (23). Complete open-reading frame sequences for all constructs were confirmed by automated dideoxy sequencing in the University of Utah Health Sciences DNA Sequencing Core Facility. Primer sequences used to generate constructs are provided in the supplemental material.

Cell lines and culture

Cos7L cells were maintained in Dulbecco's modified Eagle's medium (DMEM) supplemented with 10% FBS, GlutaMAX, and pen-strep. CCRF-CEM cells were maintained in RPMI-1640 media supplemented with 10% FBS, GlutaMAX and pen-strep. Gal4-UAS-TK-Luc HEK 293T-Rex cells (referred to as HEK293 reporter cells throughout the text) were generously provided by Raphael

Margueron of the Institut Curie- Unité de Génétique et de Biologie du Développement, Paris, France, and maintained in DMEM supplemented with 10% FBS, GlutaMAX and pen-strep.

Luciferase assays

Luciferase assays were performed using a dual luciferase assay kit according to manufacturer protocols (Promega). Briefly, 3.25 x 10^5 Gal4-UAS-TK-Luc HEK 293T-Rex cells were seeded per well in 6-well plates. Transfection of plasmids expressing SNAG:Gal4 derivatives and additional plasmids shown in figures was performed in triplicate using Lipofectamine 2000 according to manufacturer protocols. Reporter activity was measured in cell lysates collected between 24 and 48 h posttransfection. Firefly luciferase activity was normalized to a constitutive, cotransfected Renilla luciferase control. Statistical significance was determined by two-sided unpaired t-tests (*P*-values, * <0.05, ** <0.005, *** <0.0005, **** <0.00005).

ChIP-qPCR

Briefly, transfection of Gal4-UAS-TK-Luc HEK293 T-Rex cells was performed using Lipofectamine 2000 according to manufacturer protocols. 48 h posttransfection cells were crosslinked in 1% formaldehyde for 15 min at room temperature. Crosslinking was terminated with 125mM glycine for 5 min at room temperature. Cells were washed with ice cold PBS and lysed in ice cold Farnham lysis buffer (24, 25). Cell lysates were sonicated at 5 watts for 5 s on, 25 s off

intervals for a total of 1 min and 30 s. Antibody coated Dynabeads (Thermo Fisher Scientific) were added to clarified lysates and incubated at 4°C overnight. Beads were washed three times with LiCl buffer (100mM Tris, 500mM LiCl, 1% NP-40, 1% sodium deoxycholate, pH 7.5) followed by a single wash in TE buffer (10mM Tris, 1mM EDTA, pH 8.0). Crosslinks were reversed in IP elution buffer (100mM $NaHCO_3$, 1% SDS) at 65°C overnight. Immunoprecipitated DNA was recovered using ChIP DNA Clean and Concentrator (Zymo Research) and quantified using the PicoGreen system (Thermo Fisher Scientific). PCR was performed using the Luc-TSS-FWD and Luc-TSS-REV primers (see supplemental information). Fold enrichment was determined by the $\Delta\Delta C_t$ method and normalized to input DNA. All qPCR reactions were performed in triplicate using SsoFast EvaGreen Supermix reagent (Bio-Rad).

Zebrafish morpholino and mRNA injections

Zebrafish were maintained in the aquatics shared resource at the Huntsman Cancer Institute. Morpholinos were purchased from Gene Tools, LLC. mRNAs were generated using mMessage mMachine T7 transcription reactions (Ambion, Thermo Fischer). Embryos were microinjected at the 1 cell stage with 1 nl of 20 μM splice blocking *gfi1aa* and 200 μM *kdm1a* morpholino with 10% phenol red and sterile water. mRNAs were coinjected with morpholino at a final concentration of 100 ng/μl. Unspliced *gfi1aa* mRNA was detected in *gfi1aa* morpholino injected embryos by PCR amplification of the morpholino targeting intron 1/exon 2 boundary within cDNA preparations as described previously (26). Morpholino and

primer sequences are provided in supplemental materials. Dechorionated, 48 h postfertilization (hpf) embryos were stained in a solution of 40% ethanol, 0.65% H_2O_2, 0.01M sodium acetate, 0.6mg/ml *o*-dianisidine in the dark on a rolling rotisserie for 2 h. Embryos were thoroughly washed with PBS-T and visualized on a dissecting microscope. Primitive erythropoiesis was quantified using a scoring system from 1 to 4, indicating minimal (1), modest (2), moderate (3), and complete (4) hemoglobinization as revealed by *o*-dianisidine staining. More than 60 embryos were scored for each experimental condition. Statistical significance was determined by Wilcoxon-Mann-Whitney testing in GraphPad Prism 6 (*P*-values, * <0.05, ** <0.005, *** <0.0005, **** <0.00005).

Transfection, immunoprecipitation, and drug treatments

Cos7L cells were transfected using Lipofectamine 2000 (Life Technologies) as previously described (26). Cell lysates were harvested in phospho protecting lysis buffer (PPLB) 48 h posttransfection and immunoprecipitation performed from clarified cell lysates using α-FLAG (M2) antibody and Protein G Sepharose beads (26). Immune complexes and clarified lysates were probed using antibodies indicated in the figures and figure legends. SMYD2 inhibitor, LLY-507, was prepared in DMSO and dosed as indicated (27). The LSD1 inhibitor, HCI-2509, was prepared in DMSO and dosed as indicated (28).

In vitro peptide binding

GFI1 SNAG domain peptides were synthesized by GenScript to > 95% purity and conjugated to biotin. SNAG domain peptides were incubated at 4°C with purified human recombinant LSD1 or extracts prepared from CCRF-CEM cells under nondenaturing conditions. Biotinylated peptides and bound LSD1 were isolated using streptavidin agarose beads (Thermo Fisher Scientific), washed extensively in incubation buffer (50mM HEPES, pH 7.5) and subjected to α-LSD1 western blotting as described above.

Methyltransferase assays

GFI1 SNAG domain peptides were incubated with human recombinant SMYD2 (Cayman Chemical) and ^{3}H-S-adenosylmethionine (SAM) at 30°C. Biotinylated peptide substrates were purified from reaction mixtures using streptavidin agarose beads (Thermo Fisher Scientific). Beads were washed extensively with reaction buffer, resuspended in scintillation cocktail and ^{3}H incorporation measured using a Beckman Coulter LS 6500 liquid scintillation counter. Statistical significance was determined by two-sided unpaired t-tests (*P*-values, * <0.05, ** <0.005, *** <0.0005, **** <0.00005).

shRNA and flow cytometry

GFI1 depletion was achieved in cell lines via an shRNA targeting the 3' UTR of *GFI1* subcloned into the pMKO1 vector (a generous gift of Dr. Stephen Lessnick), relative to content matched scramble control. Selection was performed

in growth medium containing 3μg/ml puromycin for 3 days. Apoptosis in stable cells was determined by annexin V/propidium iodide (PI) staining on a FACSCanto Analyzer (BD).

RNA sequencing

Transcriptome RNA sequencing was performed as follows. Total RNA from 4 biological replicates for each experimental condition was isolated using the Qiagen RNeasy Mini Kit with on column DNAse treatment. RNA quality was confirmed by RNA integrity number scoring on the Agilent RNA ScreenTape Assay. The Illumina TruSeq Stranded Total RNA Sample Prep Kit with Ribo-Zero Gold was used for sequencing library preparation. Illumina HiSeq 50 Cycle Single-Read Sequencing version 4 sequencing protocol was performed and reads were aligned to the hg38 human genome assembly. Differentially expressed genes were identified at a greater than 2 fold change threshold and an adjusted *P*-value of 0.01 relative to vector control CCRF-CEM cells.

Results

SNAG K8 methylation by SMYD2 promotes LSD1 binding and GFI1 mediated transcriptional repression

Transcriptional repression by GFI1 can be predominantly attributed to LSD1 recruitment by the SNAG domain (1). To gain additional insights into determinants of LSD1 binding, we examined the primary structure of GFI family proteins to find near complete conservation of a -^{8}KSKK^{11}- motif embedded within the SNAG

domains of GFI1 and GFI1B among diverse species (Figure 3.1). An analogous motif in p53 is subject to methylation by specific methyltransferases to dynamically regulate p53 function. We hypothesized methylation on -^{8}KSKK^{11}- lysine residues may similarly impact functions of the GFI1 SNAG domain.

To test the contribution of -^{8}KSKK^{11}- motif residues to GFI1 mediated transcriptional repression we utilized a cell based, heterologous transcriptional reporter stably integrated into the host cell genome and regulated by a concatemer of five Gal4–UAS elements juxtaposed to an HSV thymidine kinase (TK) minimal promoter (Figure 3.1) (29). This system enables transcriptional output to be attributed specifically to motifs in GFI1 incorporated into fusion proteins with Gal4, free from partnerships GFI1 might have with other DNA binding proteins on an endogenous promoter. Likewise, the system permits transcriptional output to be measured in a chromatinized context and for assessment of promoter occupancy and histone modifications to be made in a unified fashion. When fused to the Gal4 DNA binding domain, wild type (WT) SNAG or derivatives with discrete mutations can be examined for impact upon transcription, LSD1 recruitment and H3K4 methylation status. We generated WT, K8L, K10L, and K11L SNAG:Gal4 fusion proteins (Figure 3.1) and tested their impact on transcriptional output from the integrated reporter. The SNAG domain alone, fused to Gal4, was sufficient to repress reporter output comparable to a GFI1-WT:Gal4 fusion protein lacking zinc fingers 4-6 (Figure 3.1 and data not shown), confirming a dominant role for the SNAG domain in GFI1 mediated transcriptional repression. Leucine substitution at K8 impaired repressor function, quantitatively similar to that observed with P2A

substitution (Figure 3.1) (2). Leucine substitution at either K10 or K11 had little impact on repressor function (Figure 3.1).

LSD1 is the dominant effector of GFI1 mediated transcriptional repression and directly binds the GFI1 SNAG domain (1). To test whether the transcriptional repression defect of the K8L derivative reflected impaired LSD1 binding, we expressed FLAG-tagged WT, K8L, P2A, and ΔSNAG derivatives of GFI1 in Cos7L cells and determined their ability to engage endogenously expressed LSD1 in α-FLAG coprecipitation assays. K8L substitution abolished the interaction between GFI1 and LSD1, comparable to SNAG domain deletion or P2A substitution (Figure 3.1). These results indicate the GFI1 SNAG domain alone is sufficient for transcriptional repression, and by supporting LSD1 recruitment, K8 is essential to this function of the SNAG domain.

We then tested whether methyltransferases, SMYD2, SETD7, and G9a influence SNAG domain mediated transcriptional repression. Methyltransferases were coexpressed with SNAG-WT:Gal4 and reporter activity quantified. SMYD2 coexpression enhanced repression by SNAG-WT:Gal4, while SETD7 and G9a coexpression had minimal impact (Figure 3.2). Given the functional relationship between SMYD2 and transcriptional repression, we asked whether SMYD2 bound the SNAG domain. FLAG-tagged SMYD2 was transiently expressed alongside Gal4-tagged WT and K8L SNAG domains. Gal4 western blot of FLAG-immune complexes revealed an equivalent interaction between SMYD2 and both WT and K8L SNAG domains (Figure 3.2). We then tested SMYD2 methyltransferase activity towards GFI1 SNAG peptide either without or with the K8L substitution

(Figure 3.2). WT or K8L SNAG peptides were incubated with recombinant SMYD2 and ^{3}H-SAM. SNAG peptides were purified via their biotin tag and methylation measured by scintillation counting. K8 substitution significantly reduced activity of SMYD2 toward the SNAG domain, suggesting K8 is a SMYD2 methyl acceptor residue in GFI1. To test whether K8 is required for enhanced repressor function conferred by SMYD2, we expressed WT or K8L SNAG:Gal4 alongside SMYD2 using the integrated reporter system. SNAG-K8L:Gal4 was insensitive to additional repression by SMYD2 (Figure 3.2). Likewise, an inhibitor of SMYD2, LLY-507, caused concentration dependent reversal of repression only for SNAG-WT but not for its K8L variant (Figure 3.2 and data not shown) (27). These data suggest SMYD2 contributes to GFI1 mediated transcriptional repression via methylation of K8 within the SNAG domain.

SNAG domain K8 methylation recruits LSD1 for H3K4 demethylation and gene repression

SMYD2 binds and methylates the GFI1 SNAG domain at K8 and contributes to GFI1 repressor function. To address directly the role of K8 methylation for SNAG—LSD1 binding, we synthesized biotinylated SNAG peptide dimethylated on K8 (SNAG-K8me2) then compared its interaction with LSD1 to that of unmethylated SNAG peptide. Recombinant human LSD1 (hLSD1) and CCRF-CEM cell lysates containing LSD1 were incubated with biotinylated SNAG or SNAG-K8me2 peptides, then copurified on streptavidin conjugated beads. SNAG-K8me2 peptide extracts hLSD1 from solution quantitatively while unmethylated

SNAG peptide bound only a small minority of the hLSD1 available, leaving >90% in solution (Figure 3.3). Similarly, the SNAG—LSD1 interaction was nearly undetectable using unmethylated SNAG peptide and CCRF-CEM extracts, while LSD1 binding increased dramatically using SNAG-K8me2 peptide (Figure 3.3). These findings indicate that K8 methylation strongly favors SNAG—LSD1 binding.

To clarify LSD1's contribution to GFI1 transcriptional repression in the integrated reporter system, LSD1 was coexpressed alongside SNAG-WT:Gal4 and reporter activity quantified. Enforced LSD1 expression enhances transcriptional repression by SNAG-WT:Gal4 with no discernible impact on SNAG-K8L:Gal4 (Figure 3.4 and data not shown). Expression of a catalytically inactive LSD1 derivative (K661A) reverses the enhanced repression conferred by WT LSD1 expression (Figure 3.4) while an LSD1 inhibitor, HCI-2509, causes concentration dependent reversal of repression by SNAG-WT:Gal4 but not SNAG-K8L:Gal4 (Figure 3.4 and data not shown) (28). We then used chromatin immune precipitation and quantitative PCR (ChIP-qPCR) to investigate how loss of K8 methylation might impact LSD1 recruitment and H3K4me2 status, and to correlate with transcriptional output from the integrated reporter system. ChIP-qPCR revealed WT and K8L SNAG:Gal4 equivalently bind the reporter promoter (Figure 3.4). However, K8L substitution profoundly reduced LSD1 occupancy compared to SNAG-WT:Gal4 (Figure 3.4) and resulted in enrichment of H3K4me2 marks at the promoter locus (Figure 3.4). Combined, these data indicate K8 methylation strongly favors LSD1 recruitment by the SNAG domain to direct H3K4 demethylation and to bring about transcriptional repression.

K8 methylation contributes to GFI1 function

during developmental erythropoiesis

GFI1 contributes to early myelo and lymphopoiesis in mice and humans (3). In the developing zebrafish, the GFI1 homolog, *gfi1aa*, is required for primitive erythropoiesis (30). We leveraged this functional requirement to test the importance of GFI1 K8 methylation *in vivo*. Hemoglobinized primitive erythrocytes in zebrafish were visualized by whole mount *o*-dianisidine staining at 48 h postfertilization (Figure 3.5). Embryo staining was scored across a numeric continuum representing minimal to complete erythropoiesis (26). Consistent with previous reports, morpholino mediated depletion of *gfi1aa* impairs primitive erythropoiesis and this defect could be complemented by coinjection of morpholino resistant mRNA encoding WT rat *Gfi1* (Figure 3.5) (26, 30). However, coinjection of *Gfi1* K8L mRNA failed to complement erythropoiesis defects brought on *gfi1aa* depletion, just as was observed using mRNA for the LSD1 binding deficient mutant, *Gfi1* P2A. Notably, morpholino mediated depletion of the zebrafish LSD1 homolog, *kdm1a*, phenocopies the defect in primitive erythropoiesis shown for *gfi1aa* depletion. This phenotype was reversed by coinjection of wild type human LSD1 mRNA (Figure 3.5). These data highlight the critical importance K8 dependent GFI1—LSD1 binding to a GFI1 driven outcome in developmental hematopoiesis and suggest K8 methylation is a critical determinant of GFI1—LSD1 axis function.

K8 methylation is required for GFI1 progrowth and survival functions in lymphoid leukemia cells

GFI1 is critical for the initiation and maintenance of lymphoblastic leukemia (11). GFI1 conditional deletion impedes lymphoma development in animal models and causes regression of established lymphomas (12, 31). We tested the importance of K8 for progrowth and survival functions of GFI1 in lymphoid leukemia cells. Using a *GFI1* targeted shRNA, we depleted GFI1 in the lymphoblastic leukemia cell line, CCRF-CEM (Figure 3.6) and determined apoptosis relative to a content matched vector control (Figure 3.6). Apoptosis was determined by flow cytometry using concurrent propidium iodide (PI) and annexin V staining. Control cells display size and complexity features consistent with lymphocytes. A limited subset of cells underwent apoptosis (27% annexin V+/PI-). Following GFI1 depletion, the lymphocyte population displayed increased apoptosis (62% annexin V+/PI-) and a population with small size and widely divergent side scatter emerged. This emergent population, which increased with time following GFI1 depletion, is consistent with secondary necrosis (98% annexin V+/PI+) (Figure 3.6). Quantifying both apoptotic and secondarily necrotic populations reveals that GFI1 is required for CCRF-CEM cell survival (Figure 3.6).

To determine the impact of K8 methylation and LSD1 recruitment on cell survival, we determined cell viability in stable CCRF-CEM cell lines expressing FLAG-tagged WT, P2A, and K8L GFI1 derivatives or vector control and simultaneously depleted of endogenous GFI1. We confirmed shRNA mediated depletion of endogenous GFI1 and enforced expression of FLAG tagged GFI1

derivatives by western blot (Figure 3.7). Cell lines depleted of endogenous GFI1 then reexpressing wild type GFI1 retained growth potential, while those expressing P2A, K8L, or vector control failed to expand (Figure 3.7). These data indicate GFI1 derivatives with defective LSD1 binding cannot functionally compensate for GFI1 loss and suggest that K8 methylation may be a critical determinant of lymphoid leukemia cell viability.

Total RNA from these cell lines was subjected to RNA sequencing. Clustering provided a list of genes upregulated with endogenous GFI1 depletion alone and in GFI1 depleted cells expressing WT, K8L, and P2A GFI1 derivatives (Figure 3.7). Genes upregulated following endogenous GFI1 depletion alone, as well as in GFI1 depleted cells expressing K8L and P2A GFI1 derivatives, but not WT GFI1 were subjected to gene ontology (GO) functional term enrichment (Figure 3.7). This list was enriched for genes associated with programmed cell death/regulation of apoptosis and for loss of transcriptional repression (Figure 3.7). These results suggest K8 methylation is a critical determinant of LSD1 recruitment by GFI1 to support GFI1 progrowth and survival functions in lymphoblastic leukemia cells.

Discussion

GFI1 is a transcriptional repressor and master regulator of cell fate, differentiation and survival in normal and malignant hematopoiesis (1). The N-terminal SNAG domain is the dominant repressor element within GFI1. Yet, a detailed understanding of mechanisms governing transcriptional control by GFI1

is lacking (1, 32). We focused on the SNAG domain, and a -^{8}KSKK^{11}- motif embedded within the SNAG domain to gain additional mechanistic insights into GFI1 mediated transcriptional repression. We find SMYD2 methylates K8 within the SNAG domain, that K8 methylation favors LSD1 recruitment and that this SNAG—LSD1 relationship is required for GFI1—LSD1 axis functions in transcriptional repression and hematopoietic cell fate determination.

Nonhistone proteins, including transcription factors, are subject to lysine methylation. Among them are p53, RB, E2F1, STAT3, MYOD, RELA/p65, AR, ERα, and CEBPβ (33-35). These factors play critical regulatory functions in cell growth and survival, and likewise instruct the establishment and maintenance of the differentiated state. Thus, lysine methylation occupies a central nexus for cell fate determination, making its coordinate control mechanisms among transcription factors essential for normal homeostasis. A growing roster of methyltransferases is responsible for lysine methylation among nonhistone proteins and transcriptional regulators, including SETD7, G9a, NSD2, SETD6, SMYD2, and SMYD3 (33-35). Often the methylation events catalyzed by these enzymes can be reversed by the methyl lysine specific demethylase, LSD1 (33, 35). The mechanistic and molecular consequences of lysine methylation vary for each transcription factor, and perhaps for each methylation event and stoichiometry. For example, RB is methylated at K860 by SMYD2, promoting its interaction with Polycomb group repressor protein L3MBTL1 and correlating with cell cycle arrest (36). However, RB can also be methylated at K873 by SETD7, which promotes RB's interaction with the heterochromatin protein, HP1, and supports transcriptional repression, cell cycle

arrest, and myogenic differentiation (37). These examples demonstrate discrete methylation events that alter protein—protein and transcriptional coregulator interactions to carry out specific transcriptional and biological outcomes. This same principle is reflected in our description of the SMYD2—GFI1—LSD1 functional axis. SMYD2 methylates K8 to favor LSD1 binding. K8L substitution abolishes binding between GFI1 and LSD1, profoundly impairs transcriptional repression by the SNAG domain and renders GFI1 inactive in hematopoietic complementation and growth control. Yet, neither K10L nor K11L substitutions impact these phenotypes, suggesting a measure of residue selectivity in controlling GFI1 function. Further studies will be needed to determine if K8 methylation within the -^{8}KSKK^{11}- motif of GFI1's SNAG domain alone governs LSD1 recruitment, or if instead, it is one of many methyl group modifications via distinct methyltransferases that are integrated to modulate transcriptional repression. Given that SMYD2 is a common regulator for GFI1 and p53, and that GFI1 exerts dominion over p53 DNA damage and proapoptotic responses, it is attractive to speculate that other p53 -^{370}KSKK^{373}- methyltransferases might modify analogous lysine residues in the GFI1 -^{8}KSKK^{11}- motif to achieve coordinate control over their counter regulatory relationship. Functional grouping of methyltransferases and demethylases toward transcription factor pairs has obvious advantages when minimizing regulatory complexity is a goal. We are actively pursuing this notion of shared regulatory factors for GFI1 and its transcriptional partners.

Our findings add to an emerging choreography of posttranslational modifications that regulate GFI1 function and provide an example for how

coincident regulatory inputs may restrict the activity of an epigenetic effector like LSD1. We recently described a role for GFI1 SUMOylation in recruiting the LSD1/CoREST complex to GFI1 regulated promoters (26). We envisioned LSD1 recruitment via the GFI1 SNAG domain could be stabilized, and its demethylase activity stimulated by interactions between CoREST and SUMOylated GFI1. Thus, SUMOylation provides one regulatory input for LSD1 activity when recruited to GFI1 regulated genes via the SNAG domain. Our finding that K8 methylation strongly favors, and K8L substitution abolishes LSD1 binding, GFI1 mediated transcriptional repression and cell fate decisions suggests a second regulatory input for the GFI1—LSD1 axis. Our data support a working model for transcriptional control via the SMYD2—GFI1—LSD1 axis (Figure 3.8) where transcriptional repression by GFI1 is favored by concurrent, SMYD2 mediated K8 methylation and SUMOylation within the GFI1 linker, together facilitating LSD1/CoREST recruitment and enabling CoREST dependent activation of LSD1 demethylase activity at target genes (26). Future work will clarify how these and other posttranslational modifications regulate GFI1 transcriptional control to govern cell fate and survival decisions in normal and malignant hematopoiesis.

References

1. **Moroy T, Vassen L, Wilkes B, Khandanpour C.** 2015. From cytopenia to leukemia: the role of Gfi1 and Gfi1b in blood formation. Blood **126:**2561.

2. **Grimes HL, Chan TO, Zweidler-McKay PA, Tong B, Tsichlis PN.** 1996. The Gfi-1 proto-oncoprotein contains a novel transcriptional repressor domain, SNAG, and inhibits G1 arrest induced by interleukin-2 withdrawal. Mol Cell Bio **16:**6263.

3. **van der Meer LT, Jansen JH, van der Reijden BA.** 2010. Gfi1 and Gfi1b: key regulators of hematopoiesis. Leukemia **24:**1834–1843.

4. **Fiolka K, Hertzano R, Vassen L, Zeng H, Hermesh O, Avraham KB, Dührsen U, Möröy T.** 2006. Gfi1 and Gfi1b act equivalently in haematopoiesis, but have distinct, non-overlapping functions in inner ear development. EMBO Rep **7:**326–333.

5. **Zeng H, Yücel R, Kosan C, Klein-Hitpass L, Möröy T.** 2004. Transcription factor Gfi1 regulates self-renewal and engraftment of hematopoietic stem cells. The EMBO Journal **23:**4116–4125.

6. **Hock H, Hamblen MJ, Rooke HM, Schindler JW, Saleque S, Fujiwara Y, Orkin SH.** 2004. Gfi-1 restricts proliferation and preserves functional integrity of haematopoietic stem cells. Nature **431:**1002–1007.

7. **Yücel R, Karsunky H, Klein-Hitpass L, Möröy T.** 2003. The transcriptional repressor Gfi1 affects development of early, uncommitted c-Kit+ T cell progenitors and CD4/CD8 lineage decision in the thymus. J Exp Med **197:**831–844.

8. **Xia J, Bolyard AA, Rodger E, Stein S, Aprikyan AA, Dale DC, Link DC.** 2009. Prevalence of mutations in ELANE, GFI1, HAX1, SBDS, WAS and G6PC3 in patients with severe congenital neutropenia. Br J Haematol **147:**535–542.

9. **Person RE, Li F-Q, Duan Z, Benson KF, Wechsler J, Papadaki HA, Eliopoulos G, Kaufman C, Bertolone SJ, Nakamoto B, Papayannopoulou T, Grimes HL, Horwitz M.** 2003. Mutations in proto-oncogene GFI1 cause human neutropenia and target ELA2. Nat Genet **34:**308–312.

10. **Phelan JD, Shroyer NF, Cook T, Gebelein B, Grimes HL.** 2010. Gfi1-cells and circuits: unraveling transcriptional networks of development and disease. Current Opinion in Hematology **17:**300–307.

11. **Khandanpour C, Möröy T.** 2013. Gfi1 as a regulator of p53 and a therapeutic target for ALL. Oncotarget **4:**374.

12. **Khandanpour C, Phelan JD, Vassen L, Schütte J, Chen R, Horman SR, Gaudreau M-C, Krongold J, Zhu J, Paul WE, Dührsen U, Göttgens B, Grimes HL, Möröy T.** 2013. Growth factor independence 1 antagonizes a p53-induced DNA damage response pathway in lymphoblastic leukemia. Cancer Cell **23:**200–214.

13. **McGhee L, Bryan J, Elliott L, Grimes HL, Kazanjian A, Davis JN, Meyers S.** 2003. Gfi-1 attaches to the nuclear matrix, associates with ETO (MTG8) and histone deacetylase proteins, and represses transcription using a TSA-sensitive mechanism. J Cell Bio **89:**1005–1018.

14. **Saleque S, Kim J, Rooke HM, Orkin SH.** 2007. Epigenetic regulation of hematopoietic differentiation by Gfi-1 and Gfi-1b is mediated by the cofactors CoREST and LSD1. Molecular Cell **27:**562–572.

15. **Nakazawa Y, Suzuki M, Manabe N, Yamada T, Kihara-Negishi F, Sakurai T, Tenen DG, Iwama A, Mochizuki M, Oikawa T.** 2007. Cooperative interaction between ETS1 and GFI1 transcription factors in the repression of Bax gene expression. Oncogene **26:**3541–3550.

16. **Laurent B, Randrianarison-Huetz V, Frisan E, Andrieu-Soler C, Soler E, Fontenay M, Dusanter-Fourt I, Duménil D.** 2012. A short Gfi-1B isoform controls erythroid differentiation by recruiting the LSD1-CoREST complex through the dimethylation of its SNAG domain. Journal of Cell Science **125:**993–1002.

17. **Kurash JK, Lei H, Shen Q, Marston WL, Granda BW, Fan H, Wall D, Li E, Gaudet F.** 2008. Methylation of p53 by Set7/9 mediates p53 acetylation and activity in vivo. Molecular Cell **29:**392–400.

18. **Huang J, Dorsey J, Chuikov S, Perez-Burgos L, Zhang X, Jenuwein T, Reinberg D, Berger SL.** 2010. G9a and Glp methylate lysine 373 in the tumor suppressor p53. J Biol Chem **285:**9636–9641.

19. **West LE, Gozani O.** 2011. Regulation of p53 function by lysine methylation. Epigenomics **3:**361–369.

20. **Huang J, Perez-Burgos L, Placek BJ, Sengupta R, Richter M, Dorsey JA, Kubicek S, Opravil S, Jenuwein T, Berger SL.** 2006. Repression of p53 activity by Smyd2-mediated methylation. Nature **444:**629–632.

21. **Theisen ER, Gajiwala S, Bearss J, Sorna V, Sharma S, Janat-Amsbury M.** 2014. Reversible inhibition of lysine specific demethylase 1 is a novel

anti-tumor strategy for poorly differentiated endometrial carcinoma. BMC Cancer **14:**752.

22. **Vassen L, Fiolka K, Mahlmann S, Möröy T.** 2005. Direct transcriptional repression of the genes encoding the zinc-finger proteins Gfi1b and Gfi1 by Gfi1b. Nucleic Acids Res **33:**987–998.

23. **Sankar S, Theisen ER, Bearss J, Mulvihill T, Hoffman LM, Sorna V, Beckerle MC, Sharma S, Lessnick SL**. 2014. Reversible LSD1 inhibition interferes with global EWS/ETS transcriptional activity and impedes Ewing sarcoma tumor growth. Clin Cancer Res **20:**4584–4597.

24. **Duan Z, Horwitz M.** 2003. Targets of the transcriptional repressor oncoprotein Gfi-1. Proc Natl Acad Sci USA **100:**5932–5937.

25. **Landt SG, Marinov GK, Kundaje A, Kheradpour P, Pauli F, Batzoglou S, Bernstein BE, Bickel P, Brown JB, Cayting P, Chen Y, DeSalvo G, Epstein C, Fisher-Aylor KI, Euskirchen G, Gerstein M, Gertz J, Hartemink AJ, Hoffman MM, Iyer VR, Jung YL, Karmakar S, Kellis M, Kharchenko PV, Li Q, Liu T, Liu XS, Ma L, Milosavljevic A, Myers RM, Park PJ, Pazin MJ, Perry MD, Raha D, Reddy TE, Rozowsky J, Shoresh N, Sidow A, Slattery M, Stamatoyannopoulos JA, Tolstorukov MY, White KP, Xi S, Farnham PJ, Lieb JD, Wold BJ, Snyder M.** 2012. ChIP-seq guidelines and practices of the ENCODE and modENCODE consortia. Genome Research **22:**1813–1831.

26. **Andrade D, Velinder M, Singer J, Maese L, Bareyan D, Nguyen H, Chandrasekharan MB, Lucente H, McClellan D, Jones D, Sharma S, Liu F, Engel ME.** 2016. Sumoylation regulates growth factor independence (GFI)-1 in transcriptional control and hematopoiesis. Mol Cell Bio. **36:**10 1438-1450

27. **Nguyen H, Allali-Hassani A, Antonysamy S, Chang S, Chen LH, Curtis C, Emtage S, Fan L, Gheyi T, Li F, Liu S, Martin JR, Mendel D, Olsen JB, Pelletier L, Shatseva T, Wu S, Zhang FF, Arrowsmith CH, Brown PJ, Campbell RM, Garcia BA, Barsyte-Lovejoy D, Mader M, Vedadi M.** 2015. LLY-507, a Cell-active, Potent, and Selective Inhibitor of Protein-lysine Methyltransferase SMYD2. J Biol Chem **290:**13641–13653.

28. **Sorna V, Theisen ER, Stephens B, Warner SL, Bearss DJ, Vankayalapati H, Sharma S.** 2013. High-Throughput Virtual Screening Identifies Novel N ′-(1-Phenylethylidene)-benzohydrazides as Potent, Specific, and Reversible LSD1 Inhibitors. Journal of Medicinal Chemistry **56:**9496.

29. **Mozzetta C, Pontis J, Fritsch L, Robin P, Portoso M, Proux C, Margueron R, Ait-Si-Ali S.** 2014. The histone H3 lysine 9

methyltransferases G9a and GLP regulate polycomb repressive complex 2-mediated gene silencing. Molecular Cell **53**:277–289.

30. **Cooney JD, Hildick-Smith GJ, Shafizadeh E, McBride PF, Carroll KJ, Anderson H, Shaw GC, Tamplin OJ, Branco DS, Dalton AJ, Shah DI, Wong C, Gallagher PG, Zon LI, North TE, Paw BH.** 2012. Teleost growth factor independence (gfi) genes differentially regulate successive waves of hematopoiesis. Developmental Biology **373:**431–441.

31. **Phelan JD, Saba I, Zeng H, Kosan C, Messer MS, Olsson HA, Fraszczak J, Hildeman DA, Aronow BJ, Möröy T, Grimes HL.** 2013. Growth factor independent-1 maintains Notch1-dependent transcriptional programming of lymphoid precursors. PLoS Genet **9:**e1003713.

32. **Chiang C, Ayyanathan K.** 2012. Snail/Gfi-1 (SNAG) family zinc finger proteins in transcription regulation, chromatin dynamics, cell signaling, development, and disease. Cytokine & Growth Factor Reviews **24:**123–131.

33. **Biggar KK, Li SS-C.** 2014. Non-histone protein methylation as a regulator of cellular signalling and function. Nat Rev Mol Cell Biol **16:**5–17.

34. **Hamamoto R, Saloura V, Nakamura Y.** 2015. Critical roles of non-histone protein lysine methylation in human tumorigenesis. Nat Rev Cancer **15:**110–124.

35. **Stark GR, Wang Y, Lu T.** 2010. Lysine methylation of promoter-bound transcription factors and relevance to cancer. Cell Research **21:**375–380.

36. **Saddic LA, West LE, Aslanian A, Yates JR, Rubin SM, Gozani O, Sage J.** 2010. Methylation of the retinoblastoma tumor suppressor by SMYD2. J Biol Chem **285:**37733–37740.

37. **Munro S, Khaire N, Inche A, Carr S, La Thangue NB.** 2010. Lysine methylation regulates the pRb tumour suppressor protein. Oncogene **29:**2357–2367.

Figure 3.1. Lysine 8 contributes to transcriptional repression by the GFI1 SNAG domain. A. Primary protein sequence conservation of the GFI1 and GFI1B SNAG domains across species, compared to human GFI1. The conserved KSKK motif between residues 8-11 is shown. Sequence around the KSKK motif in human p53 is shown for comparison. B. Graphical depiction of HEK293 T-REx heterologous transcriptional reporter. A stably integrated luciferase reporter is expressed under the control of a thymidine kinase (TK) minimal promoter immediately downstream of five Gal4-UAS elements. Structure of fusion proteins, comprised of wild type (WT) or K→L derivative SNAG domains (as shown) and the Gal4 DNA binding domain, residues 1-147, are shown. C. Impact of SNAG domain single amino acid substitutions on integrated luciferase reporter activity. HEK293 reporter cells were transiently transfected with SNAG:Gal4 fusion proteins. Firefly luciferase output was determined using a dual luciferase assay kit with firefly luciferase values normalized to a Renilla luciferase transfection control. Results are expressed as mean +/- SD from a representative experiment performed in triplicate. D. Impact of GFI1-K8L substitution on GFI1—LSD1 binding. Flag-tagged GFI1 forms were transiently expressed in Cos7L cells. GFI1 forms were immune purified with anti-FLAG antibody and coprecipitating, endogenously expressed LSD1 detected in immune complexes (IC) by western blot. Equivalent expression and precipitation of GFI1 forms was determined by western blot of whole cell lysates (WCL) and IC's, respectively.

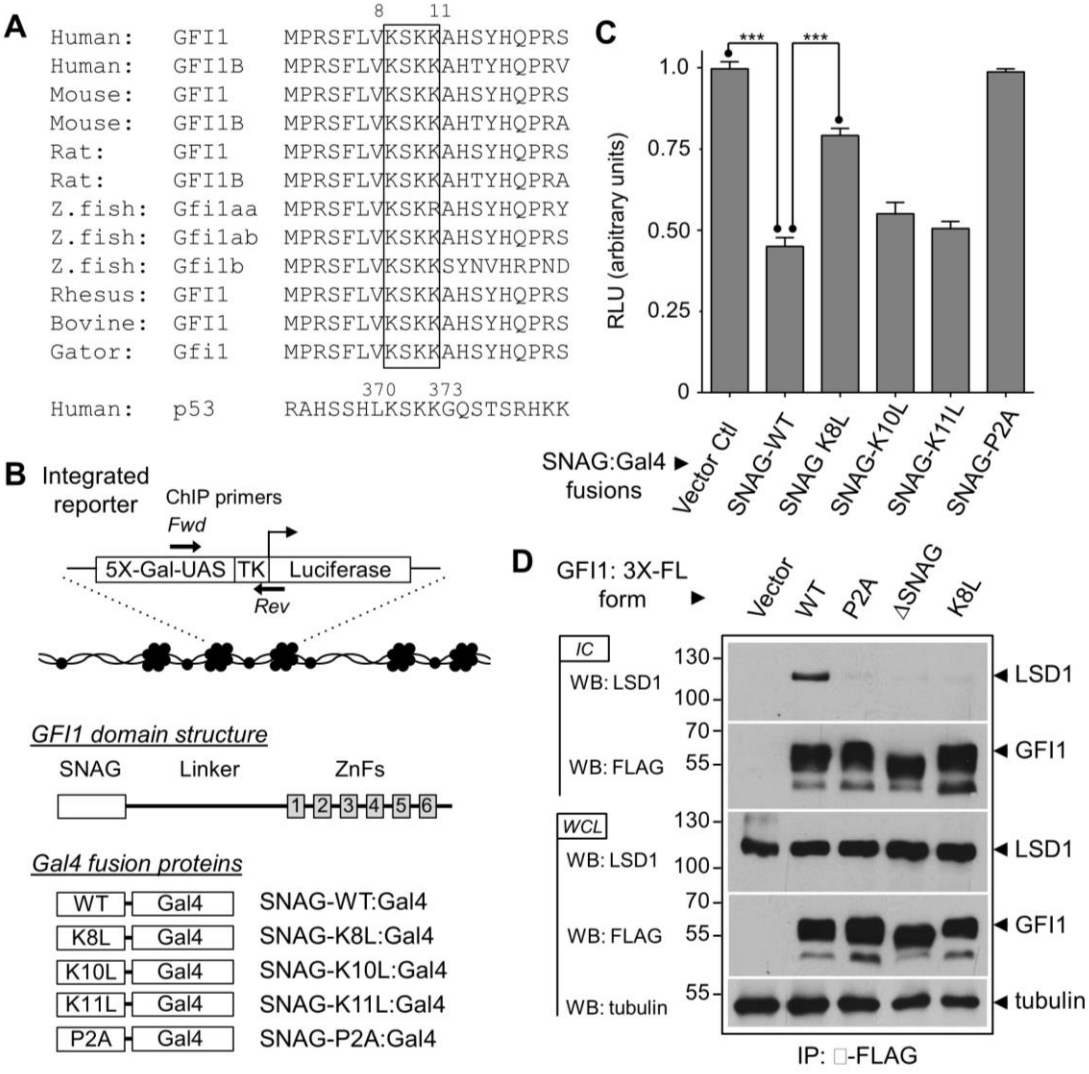

A
8 11
Human: GFI1 MPRSFLVKSKKAHSYHQPRS
Human: GFI1B MPRSFLVKSKKAHTYHQPRV
Mouse: GFI1 MPRSFLVKSKKAHSYHQPRS
Mouse: GFI1B MPRSFLVKSKKAHTYHQPRA
Rat: GFI1 MPRSFLVKSKKAHSYHQPRS
Rat: GFI1B MPRSFLVKSKKAHTYHQPRA
Z.fish: Gfi1aa MPRSFLVKSKRAHSYHQPRY
Z.fish: Gfi1ab MPRSFLVKSKKAHSYHQPRS
Z.fish: Gfi1b MPRSFLVKSKKSYNVHRPND
Rhesus: GFI1 MPRSFLVKSKKAHSYHQPRS
Bovine: GFI1 MPRSFLVKSKKAHSYHQPRS
Gator: Gfi1 MPRSFLVKSKKAHSYHQPRS
370 373
Human: p53 RAHSSHLKSKKGQSTSRHKK
B
Integrated reporter
ChIP primers
Fwd
5X-Gal-UAS
TK
Luciferase
Rev
GFI1 domain structure
SNAG
Linker
ZnFs
1 2 3 4 5 6
Gal4 fusion proteins
WT Gal4 SNAG-WT:Gal4
K8L Gal4 SNAG-K8L:Gal4
K10L Gal4 SNAG-K10L:Gal4
K11L Gal4 SNAG-K11L:Gal4
P2A Gal4 SNAG-P2A:Gal4
C
RLU (arbitrary units)
1.0
0.75
0.50
0.25
0
***
***
SNAG:Gal4 fusions
Vector Ctl
SNAG-WT
SNAG K8L
SNAG-K10L
SNAG-K11L
SNAG-P2A
D
GFI1: 3X-FL form
Vector
WT
P2A
ΔSNAG
K8L
IC
WB: LSD1
WB: FLAG
WCL
WB: LSD1
WB: FLAG
WB: tubulin
130
100
70
55
130
100
70
55
55
LSD1
GFI1
LSD1
GFI1
tubulin
IP: □-FLAG

Figure 3.2. SMYD2 mediated methylation at K8 of the SNAG domain contributes to repressor function. A. Impact of SMYD2, SETD7 and G9a protein lysine methyltransferase expression on transcriptional repression by the WT SNAG domain. HEK293 reporter cells were transfected with SNAG-WT:Gal4, SMYD2, SETD7 and G9a as indicated. Luciferase assays were performed as described in Figure 1B. B. SMYD2 binds the GFI1 SNAG domain. FLAG tagged SMYD2 and fusion proteins comprised of SNAG-WT or SNAG-K8L and the Gal4 DNA binding domain were expressed in Cos7L cells. FLAG-SMYD2 was immune purified from whole cell lysates (WCL) and co-precipitation of the SNAG domain was determined by Gal4 immunoblot (WB) of FLAG immune complexes (IC). A dashed vertical line indicates a lane that was removed from the film for clarity. C. SMYD2 methylates K8 of the SNAG domain. Human recombinant SMYD2 was incubated with WT and K8L SNAG domain peptides and [3H]-SAM methyl-donor. Peptide 3H incorporation was determined by liquid scintillation. D. SMYD2 augments transcriptional repression by the SNAG domain via K8. SNAG-WT:Gal4 or SNAG-K8L:Gal4 proteins were expressed independently or in conjunction with SMYD2. Luciferase output was measured as previously described. E. Impact of SMYD2 inhibition on SNAG-mediated transcriptional repression. SNAG-WT:Gal4 expressing HEK293 reporter cells were treated with SMYD2 inhibitor (LLY-507) as shown. Firefly luciferase reporter output was measured and normalized to Renilla luciferase control. For luciferase assays, results are expressed as mean +/- SD from a representative experiment performed in triplicate.

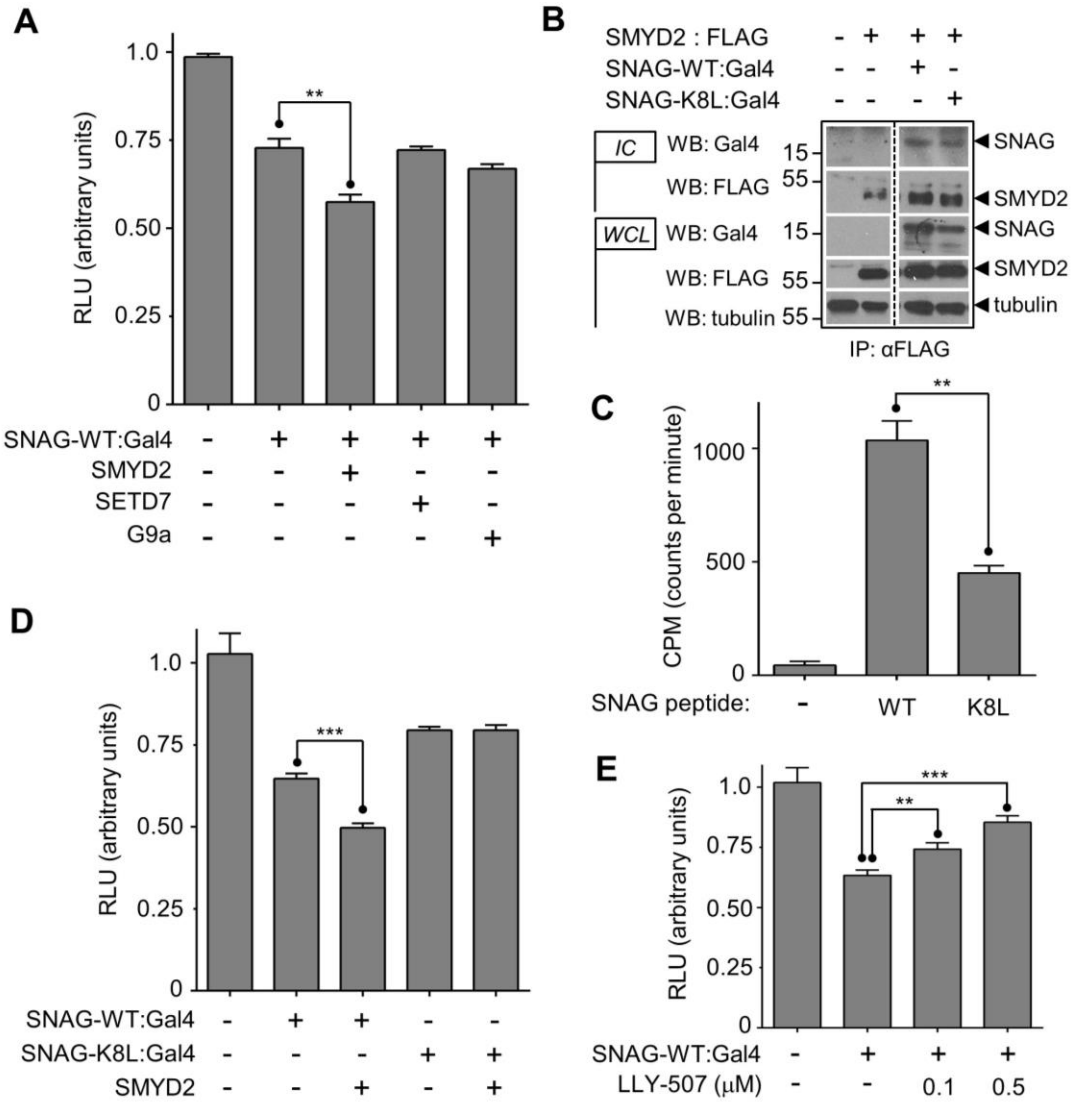

A
RLU (arbitrary units)
1.0
0.75
0.50
0.25
0
**
SNAG-WT:Gal4 - + + + +
SMYD2 - - + - -
SETD7 - - - + -
G9a - - - - +
B
SMYD2 : FLAG - + + +
SNAG-WT:Gal4 - - + -
SNAG-K8L:Gal4 - - - +
IC
WB: Gal4
WB: FLAG
WCL
WB: Gal4
WB: FLAG
WB: tubulin
15
55
15
55
55
SNAG
SMYD2
SNAG
SMYD2
tubulin
IP: αFLAG
C
CPM (counts per minute)
1000
500
0
**
SNAG peptide: - WT K8L
D
RLU (arbitrary units)
1.0
0.75
0.50
0.25
0
***
SNAG-WT:Gal4 - + + - -
SNAG-K8L:Gal4 - - - + +
SMYD2 - - + - +
E
RLU (arbitrary units)
1.0
0.75
0.50
0.25
0
***
**
SNAG-WT:Gal4 - + + +
LLY-507 (μM) - - 0.1 0.5

Figure 3.3. K8 methylation favors SNAG—LSD1 binding. A. SNAG domain K8 methylation strongly favors binding of recombinant, purified human LSD1 (hLSD1). Unmethylated (SNAG) or K8 dimethylated (SNAG-K8me2) SNAG peptides carrying a biotin moiety were incubated with hLSD1. Biotinylated SNAG peptides were purified on streptavidin-conjugated agarose beads. Beads were washed extensively in incubation buffer. LSD1 copurifying with SNAG peptides or remaining in the supernatant was detected by western blot (WB). B. K8 methylation enables SNAG—LSD1 binding from CCRF-CEM extracts. Unmethylated or K8 dimethylated SNAG domain peptides were incubated with CCRF-CEM extracts as noted. Biotinylated peptides were captured on streptavidin-conjugated agarose beads, washed and subjected to WB to detect co-purifying, endogenously expressed LSD1. LSD1 input is shown.

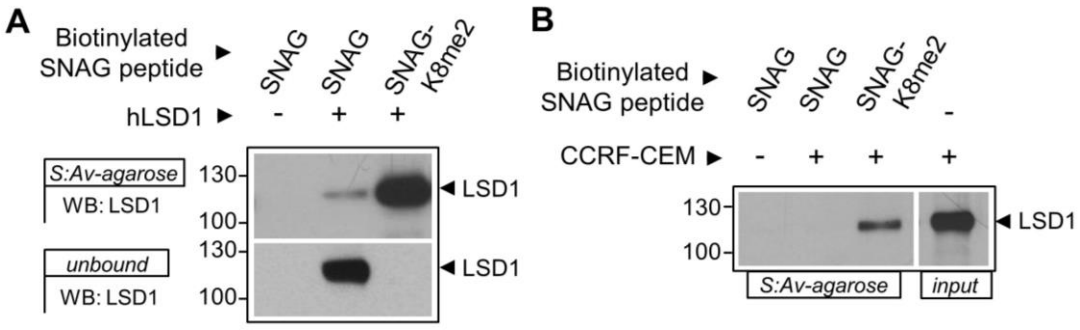

A
Biotinylated SNAG peptide ►
SNAG
SNAG
SNAG-K8me2
hLSD1 ►
-
+
+
S:Av-agarose
WB: LSD1
130
100
LSD1
unbound
WB: LSD1
130
100
LSD1
B
Biotinylated SNAG peptide ►
SNAG
SNAG
SNAG-K8me2
-
CCRF-CEM ►
-
+
+
+
130
100
LSD1
S:Av-agarose
input

Figure 3.4. K8 is required for LSD1 recruitment and H3K4 demethylation at a GFI1 target gene. A. LSD1 contributes to SNAG-mediated repression. HEK293 reporter cells were transfected with SNAG-WT:Gal4 and LSD1 as indicated. Firefly luciferase reporter output was measured, normalized to a Renilla internal control and results reported as the mean +/- SD from a representative experiment. B-C. Catalytically inactive LSD1 (K661A) (B) and LSD1 inhibition with HCI-2509 (C) antagonize SNAG-mediated transcriptional repression. HEK293 reporter cells were transfected with SNAG-WT:Gal4 and LSD1 forms or treated with HCI-2509 as shown. Reporter output was measured as in A. D-F. SNAG-K8L substitution impairs LSD1 promoter binding and stabilizes H3K4me2 marks. HEK293 reporter cells were transfected with empty vector, SNAG-WT: and SNAG-K8L:Gal4 as indicated. Cells were subjected to crosslinking, harvesting and sonication, then chromatin immunoprecipitation (ChIP) was performed using Gal4, LSD1 and H3K4me2 antibody coated beads. Quantitative PCR (qPCR) targeting the promoter of the integrated reporter (see Figure 1B) was performed using ChIP DNA as template. Fold enrichment was determined by the $\Delta\Delta C_t$ method and normalized to input DNA.

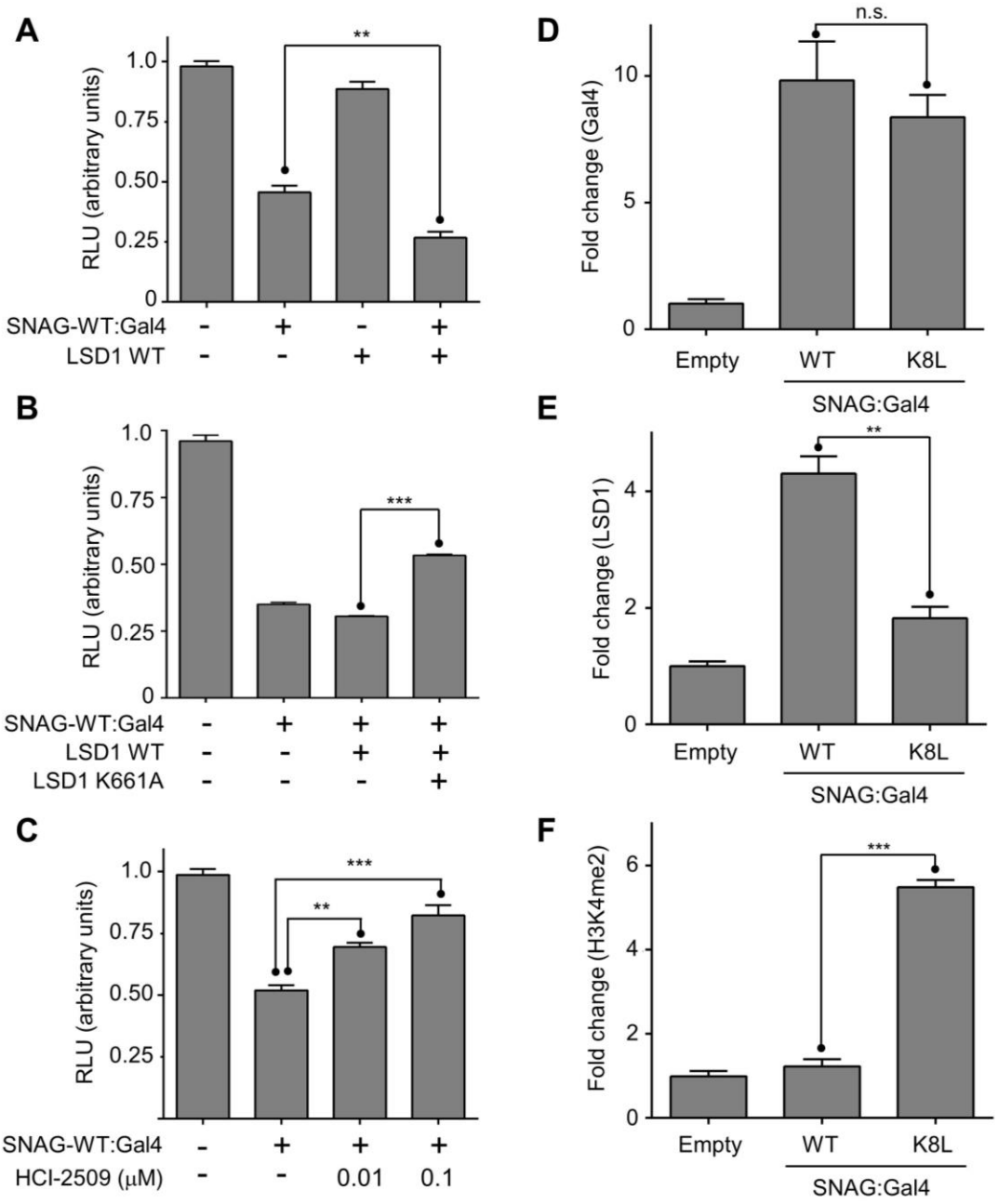

A
RLU (arbitrary units)
1.0
0.75
0.50
0.25
0
**
SNAG-WT:Gal4 - + - +
LSD1 WT - - + +
B
RLU (arbitrary units)
1.0
0.75
0.50
0.25
0
***
SNAG-WT:Gal4 - + + +
LSD1 WT - - + +
LSD1 K661A - - - +
C
RLU (arbitrary units)
1.0
0.75
0.50
0.25
***
**
SNAG-WT:Gal4 - + + +
HCI-2509 (μM) - - 0.01 0.1
D
Fold change (Gal4)
10
5
0
n.s.
Empty
WT
K8L
SNAG:Gal4
E
Fold change (LSD1)
4
2
0
**
Empty
WT
K8L
SNAG:Gal4
F
Fold change (H3K4me2)
6
4
2
0
***
Empty
WT
K8L
SNAG:Gal4

Figure 3.5. K8 methylation is required for zebrafish primitive erythropoiesis. A. Primitive erythropoiesis impaired by *gfi1aa* depletion is restored by WT GFI1 but not by K8L or P2A derivatives. Zebrafish embryos were microinjected at the 1 cell stage with a morpholino oligomer blocking *gfi1aa* pre-mRNA splicing or a content matched, scrambled sequence control as indicated (Depletion). Morpholino resistant rat *Gfi1*-WT, *Gfi1*-P2A or *Gfi1*-K8L mRNA was coinjected as indicated (Rescue). When no depletion or rescue was indicated, mock injections were performed as shown. At 48 h postfertilization (hpf) primitive erythropoiesis was revealed by *o*-dianisidine staining. Four representative embryos are shown out of more than 60 scored for each experimental condition. Primitive erythropoiesis was quantified using a hemoglobinization score from 1 to 4, indicating minimal (1), modest (2), moderate (3) and complete (4) hemoglobinization. The percentage of embryos with each respective score is reported as a stacked sum graph and statistical significance determined using a Wilcoxon-Mann-Whitney test (***, $P < 0.0005$). B. Zebrafish LSD1 homolog, *kdm1a*, depletion impairs zebrafish primitive erythropoiesis. Splice blocking *kdm1a* morpholino and morpholino resistant human WT LSD1 mRNA were injected at the 1 cell stage as indicated in the table. Primitive erythropoiesis was scored in a blinded fashion at 48hpf on the 1-4 scale as in A.

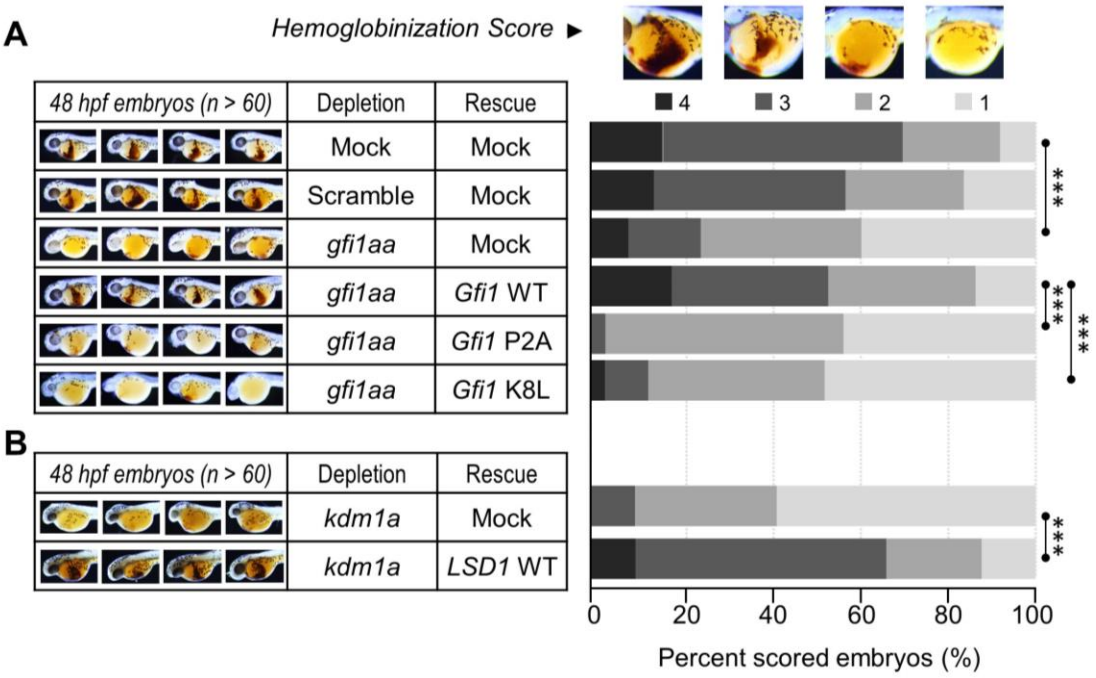

A
Hemoglobinization Score ►
48 hpf embryos (n > 60)
Depletion
Rescue
Mock
Mock
Scramble
Mock
gfi1aa
Mock
gfi1aa
Gfi1 WT
gfi1aa
Gfi1 P2A
gfi1aa
Gfi1 K8L
B
48 hpf embryos (n > 60)
Depletion
Rescue
kdm1a
Mock
kdm1a
LSD1 WT
4
3
2
1
***
***
***
***
0
20
40
60
80
100
Percent scored embryos (%)

Figure 3.6. GFI1 is required for lymphoid leukemia cell survival. A. CCRF-CEM cells are depleted of GFI1. Whole cell lysates were prepared from uninfected CCRF-CEM cells or those infected with short hairpin RNA targeting *GFI1* (shGfi1) or a content-matched scramble control (Control). GFI1 protein expression was assessed by western blot with tubulin serving as a protein loading control. B-C. GFI1 depletion leads to CCRF-CEM apoptosis and secondary necrosis. CCRF-CEM cells were transduced with scramble control vector or shGfi1 to deplete GFI1, then subjected to flow cytometry analysis. Forward scatter (FSC)/side scatter (SSC) identifies discrete cell populations. Gates were established for each population and viability assessed by annexin V and propidium iodide (PI) staining. Percentages representing viable, early (annexin V+/PI-) and late apoptosis/secondary necrosis (annexin V+/PI+) are indicated by quadrant labels and quantified as a stacked sum bar graph.

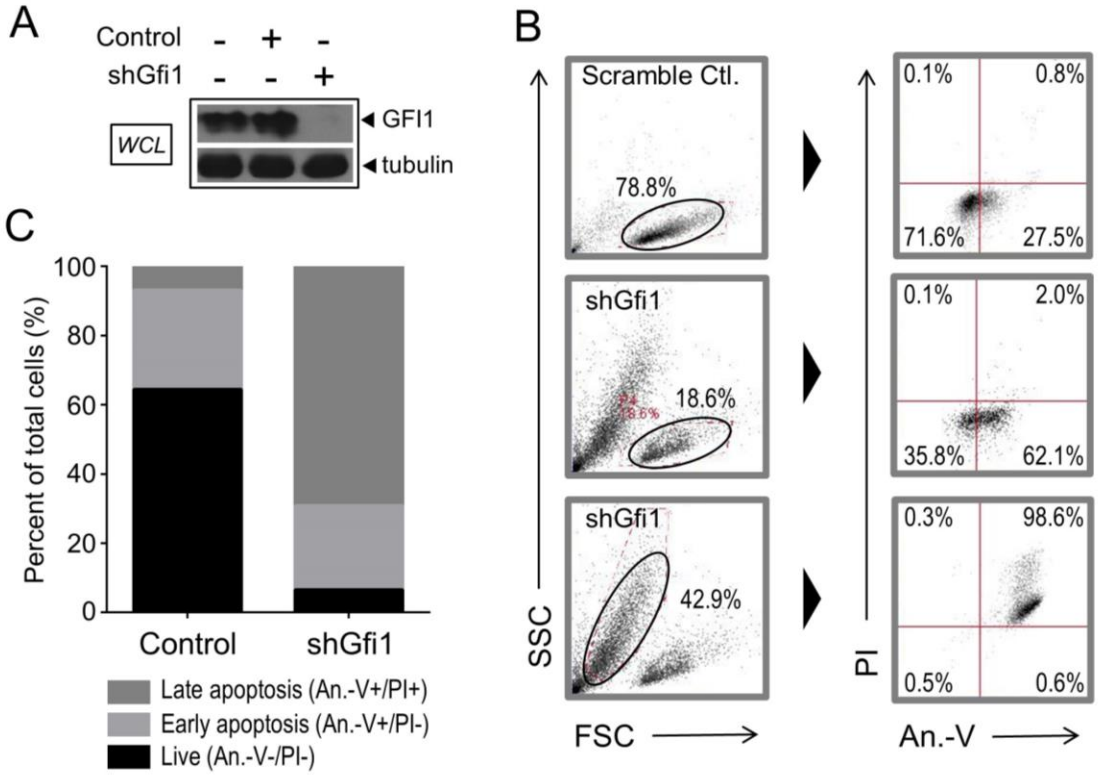

A
Control - + -
shGfi1 - - +
WCL
GFI1
tubulin
B
Scramble Ctl.
78.8%
0.1%
0.8%
71.6%
27.5%
shGfi1
18.6%
0.1%
2.0%
35.8%
62.1%
shGfi1
42.9%
0.3%
98.6%
0.5%
0.6%
SSC
FSC
PI
An.-V
C
Percent of total cells (%)
100
80
60
40
20
0
Control
shGfi1
Late apoptosis (An.-V+/PI+)
Early apoptosis (An.-V+/PI-)
Live (An.-V-/PI-)

Figure 3.7. K8 is required for GFI1 progrowth and survival functions in lymphoid leukemia cells. A. FLAG-tagged WT, P2A and K8L GFI1 expressing CCRF-CEM cell lines were depleted of endogenous GFI1 as indicated. Endogenous GFI1 and FLAG-tagged GFI1 expression was assessed by western blot, with tubulin serving as a protein loading control. B. K8 methyl defective and LSD1 binding-defective GFI1 cell lines have impaired cell growth and viability. Cell viability was determined by trypan blue exclusion counting in the indicated cell lines over 7 days following shRNA mediated GFI1 depletion. Viable cells per ml over time is reported. C. Genes associated with apoptosis regulation are upregulated in CCRF-CEM cell lines depleted of endogenous GFI1 and in GFI1 depleted cell lines expressing P2A and K8L GFI1 derivatives. Shared upregulated genes (relative to expression in naïve CCRF-CEM cells) are clustered in cells depleted of endogenous GFI1 while expressing empty vector as well as WT, P2A and K8L GFI1. 82 genes upregulated in CCRF-CEM cell lines depleted of endogenous GFI1 while expressing empty vector, P2A and K8L but not upregulated in cell lines depleted of endogenous GFI1 while expressing WT GFI1 were subjected to Gene Ontology (GO) functional term enrichment analysis. Enriched functional terms for these 82 genes are reported in the box.

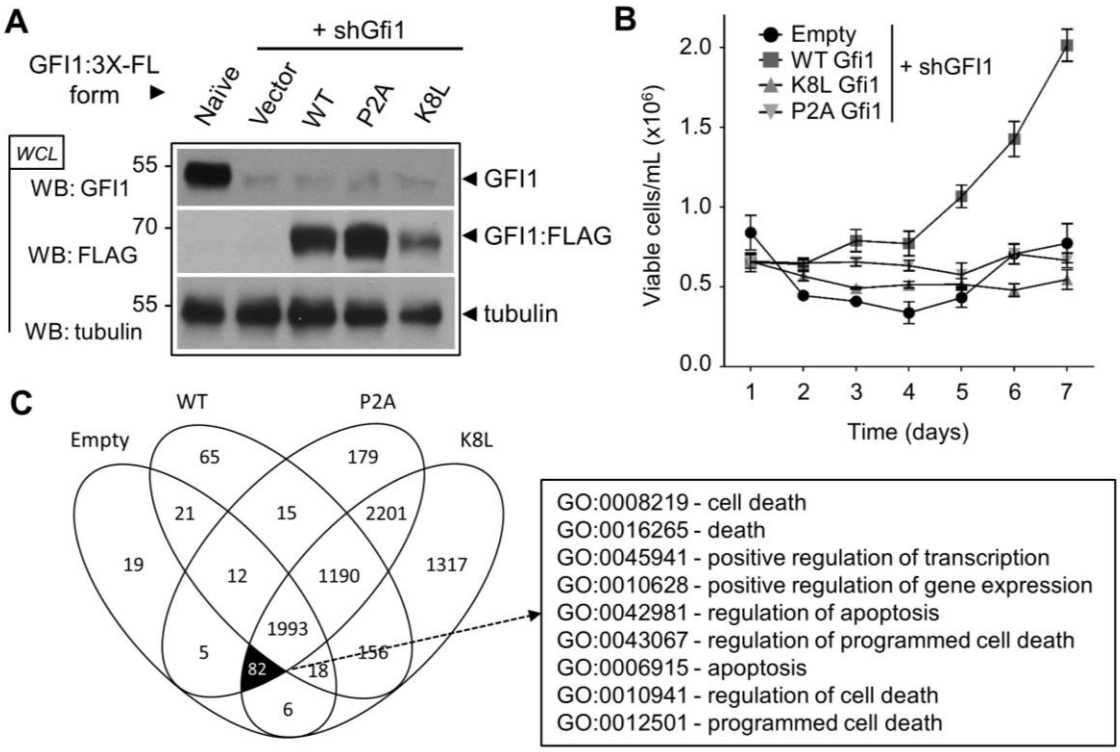
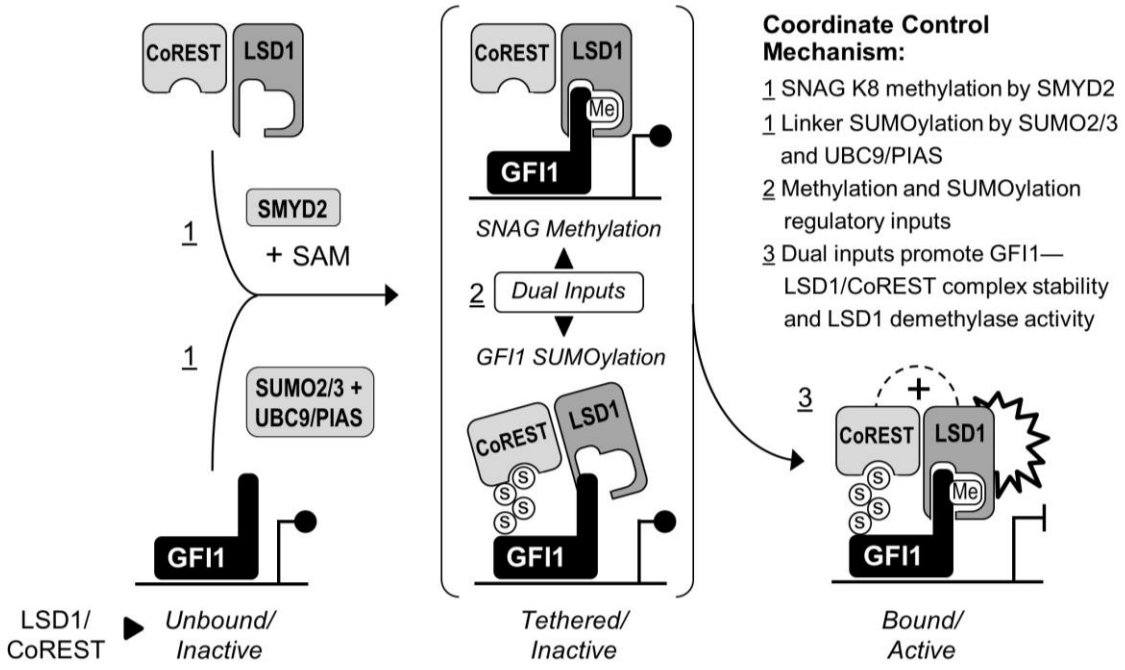

A
+ shGfi1
GFI1:3X-FL form
Naïve
Vector
WT
P2A
K8L
WCL
55
WB: GFI1
70
WB: FLAG
55
WB: tubulin
GFI1
GFI1:FLAG
tubulin
B
Empty
WT Gfi1
K8L Gfi1
P2A Gfi1
+ shGFI1
Viable cells/mL (x10^6)
2.0
1.5
1.0
0.5
0.0
1
2
3
4
5
6
7
Time (days)
C
WT
P2A
Empty
K8L
65
179
21
15
2201
19
12
1190
1317
1993
5
82
18
156
6
GO:0008219 - cell death
GO:0016265 - death
GO:0045941 - positive regulation of transcription
GO:0010628 - positive regulation of gene expression
GO:0042981 - regulation of apoptosis
GO:0043067 - regulation of programmed cell death
GO:0006915 - apoptosis
GO:0010941 - regulation of cell death
GO:0012501 - programmed cell death

Figure 3.8. A working model for GFI1 mediated LSD1/CoREST recruitment and transcriptional repression. Presented here is a model for how dual inputs to GFI1 through concurrent posttranslational modifications (methylation and SUMOylation) enable robust transcriptional repression at GFI1 regulated genes. LSD1 and CoREST are envisioned to exist in a preformed complex, but their demethylase activity would be restricted/limited by a lack of regulatory inputs. The SNAG domain has intrinsic affinity for LSD1, which is significantly enhanced by SNAG domain methylation on K8 via the action of SMYD2/S-adenosylmethionine (SAM), and perhaps altered by other KSKK modifications yet to be explored. Simultaneous SUMOylation of GFI1 facilitates an interaction with LSD1/CoREST. LSD1 demethylase activity on chromatin is activated through an interaction between SUMO2/3 assembled on GFI1 and a SUMO interacting motif in CoREST. One might see LSD1 binding to GFI1 via the SNAG domain, but without SUMOylation and CoREST dependent stimulation of LSD1 demethylase activity, its impact on transcription would be limited. Likewise, LSD1 could be brought to GFI1 regulated genes indirectly through the SUMO—CoREST interaction, but without SNAG methylation might be insufficiently stable to have a lasting impact on gene expression. This working model will inform our efforts to understand how transcriptional control by GFI1 and other SNAG transcription factors is fine tuned to the transcriptional programs they govern.


CoREST
LSD1
1
SMYD2
+ SAM
1
SUMO2/3 +
UBC9/PIAS
GFI1
LSD1/
CoREST
Unbound/
Inactive
CoREST
LSD1
Me
GFI1
SNAG Methylation
2
Dual Inputs
GFI1 SUMOylation
S
S
S
CoREST
LSD1
GFI1
Tethered/
Inactive
Coordinate Control
Mechanism:
1 SNAG K8 methylation by SMYD2
1 Linker SUMOylation by SUMO2/3
and UBC9/PIAS
2 Methylation and SUMOylation
regulatory inputs
3 Dual inputs promote GFI1—
LSD1/CoREST complex stability
and LSD1 demethylase activity
3
+
CoREST
LSD1
Me
GFI1
Bound/
Active

CHAPTER 4

CONCLUSIONS AND FUTURE DIRECTIONS

GFI1 as a model for studying mechanisms of transcriptional repression

Growth factor independence 1 (GFI1) is a transcriptional repressor and master regulator of normal and malignant hematopoiesis (1). GFI1 is comprised of an N-terminal SNAG domain, a C-terminal concatemer of six zinc fingers, and a family divergent linker region which separates them (2). GFI1 directly binds consensus sequence DNA, utilizing zinc fingers 3, 4, and 5 to do so (3). The SNAG domain is the dominant repressive element within GFI1 and individual residues within the SNAG domain are required repressor function, while others are dispensable (3).

Based on our work and the work of others, it is reasonable to consider GFI1 requiring two primary functions for transcriptional repression. First, GFI1 must be capable of binding consensus DNA sequences through zinc fingers 3-5. Disruption of zinc fingers 3, 4, or 5 abolishes repressor function and mutations within these zinc fingers give rise to neutropenia in humans (1, 4). Second, GFI1 must have an intact SNAG domain. We provide evidence here that K8 methylation is a critical determinant of SNAG domain mediated repression. We and others also demonstrate P2 is required for transcriptional repression by GFI1 (3). Additionally, we and others find R3, S4, S9, K10, and K11 are largely dispensable for repressor function (3). However, a saturating substitution based analysis of residues required for SNAG mediated repression has not been performed and other residues within the SNAG domain may also contribute to GFI1 repressor function.

The relatively simple architecture of GFI1 found in the SNAG domain and

zinc fingers makes it an ideal transcription factor to study mechanisms of transcriptional repression as such a deconstructed system is capable of revealing novel insights devoid of other complicating regulatory interactions and inputs. Furthermore, only one protein, LSD1, is known to bind the SNAG domain (1). Within our already deconstructed system we focused on a four residue KSKK motif. Yet, despite focusing on only four amino acids within a 422 amino acid protein we found a remarkable degree of regulatory complexity exists within the SNAG domain KSKK motif. We suspect this theme will be observed in future studies using deconstructed systems. Furthermore, our data provide evidence that using deconstructed biochemical systems is a powerful approach for revealing deterministic functional relationships.

Our findings here add to the emerging choreography of posttranslational modifications that regulate GFI1 function. In Chapter 2 we describe SUMOylation contributes to LSD1/CoREST binding of GFI1. In Chapter 3 we describe how methylation also contributes to LSD1 binding to the GFI1 SNAG domain. Taken together, these data suggest GFI1 methylation and SUMOylation function as part of a series of multiple regulatory inputs that modulate GFI1 function. Within this model, we propose SNAG domain methylation and linker region SUMOylation act together to facilitate LSD1/CoREST recruitment and enable CoREST dependent activation of LSD1 demethylase activity at GFI1 target genes (Figure 3.8).

Our findings here raise many additional questions regarding the functions of GFI1 as well as posttranslational modification of transcription factors generally. 1) Are methylation events within the SNAG domain KSKK motif combinatorially

regulated? And how do methyl mark combinations impact GFI1 function? 2) Is there functional cross talk between GFI1 SUMOylation and methylation? And how is the deposition of these posttranslational modifications regulated? 3) Are methylation events within the GFI1 and p53 KSKK motifs crossregulated? Are these methylation events impactful in lymphoid leukemia? 4) What is the functional breadth of transcription factor posttranslational modification in the human proteome?

Are methylation events within the SNAG domain
KSKK motif combinatorially regulated?

Our work here demonstrates SMYD2 mediated methylation at K8 in the SNAG domain is a critical determinant of GFI1 mediated transcriptional repression. We also find SETD7 methylates K10 of the SNAG domain and G9a methylates K11 of the SNAG domain (Figure 4.1). These results indicate the methyltransferase activity of SMYD2, SETD7 and G9a is conserved across the GFI1 and p53 KSKK motifs. Within p53, KSKK methylation is dynamically regulated to alter p53 function and specific methylation events are either functionally supportive or antagonistic to one another. It is currently unknown whether methylation events are similarly regulated within the SNAG domain to modulate GFI1 function.

Overcoming current *in vitro* technical limitations. In ascribing methyltransferase activities of SMYD2, SETD7 and G9a to the GFI1 SNAG domain, we have used a reductionist approach. Our experiments were performed

in vitro using the minimum required components for the enzymatic reaction, namely, a recombinant enzyme, a peptide substrate and a methyl group donor. This reductionist approach is powerful to define sufficiency of an enzyme to methylate a substrate, but also may introduce novel activities and false positives/negatives. This point is illustrated in our difficulty detecting SMYD2 mediated K8 methylation by mass spectral approaches, while we readily detect K10 and K11L methylation by SETD7 and G9a by mass spec (Figure 4.1). In future studies it may be beneficial to use full length proteins as these components are more biologically pertinent. Using full length GFI1 may also provide additional protein domains and structures not present in the SNAG domain that are required for SMYD2 mediated methylation.

Our *in vitro* approaches here also use unmethylated SNAG domain peptides, an experimental design which presupposes the methyltransferases used are capable of acting on peptides devoid of other posttranslational modifications. It is reasonable to consider these methyltransferases, and in particular, SMYD2, may require posttranslational modification at other SNAG residues for their action. Experimentally determining whether SMYD2 and other methyltransferases activities are altered by posttranslational modification at other residues remains costly and technically difficult. For example, to thoroughly characterize how methylation at each of the other KSKK motif lysines impacts SMYD2 activity at K8, peptide substrates would need to be synthesized with every possible combination of a leucine (or otherwise) substituted residue, as well as un, mono, di, and trimethylated lysine states. With five possible states at three residues, 125 (=5^3)

unique peptide combinations exist. With current peptide synthesis technologies it would be unfeasible and cost prohibitive to synthesize 125 SNAG domain peptides. Future peptide synthesis technologies may reduce cost sufficiently to warrant such an experimental approach.

SMYD2, SETD7, and G9a are all available commercially as recombinant enzymes. However, as these enzymes are also histone methyltransferases, they are typically only validated based on their activity against histone peptide substrates. This makes it difficult to differentiate true negatives from false negatives when assessing enzyme activities on SNAG peptides. This issue may also be a contributing factor to our difficulty detecting SMYD2 mediated methylation at K8 by mass spectrometry.

An alternative approach to screening each methyltransferase's activity against more than 100 unique peptide substrates would be to incubate unmodified peptides with combinations of methyltransferases. Only eight possible combinations (=2^3) exist when analyzing three methyltransferases (SMYD2, SETD7 and G9a), a more technically feasible number of combinations. If SMYD2 requires distinct methylation states at other lysine residues within the SNAG domain for its activity, such an approach would be capable of detecting this activity. In the face of complicated methylation patterns, our findings that SETD7 and G9a respectively methylate K10 and K11 in an unmodified peptide would allow us to differentiate each enzyme's function. If other confounding additional methylation events were observed in this approach, individual activities could be ascribed based on each enzymes activity on an unmodified peptide or peptides with specific

modifications. Using a limited number of modified synthesized peptides in these reactions could help clarify confounding activities as well as directly validate each enzymes activity in the presence of a specific modification. This unbiased approach allows for all possible combinations of methylation events and is likely a more technically feasible starting point for thoroughly characterizing each methyltransferase's activity in the SNAG domain.

Similar technical challenges exist for thoroughly defining the demethylase activity of LSD1 within the SNAG domain. Unlike LSD1 demethylase activity at K370 in p53, we fail to detect LSD1 demethylase activity against a K8 dimethylated SNAG domain peptide *in vitro*, relative to an H3K4 dimethylated peptide (Figure 4.2). This lack of activity may be due to a number of issues. We have yet to test whether LSD1 is capable of demethylating K8 monomethylated SNAG domain. Also, using SNAG domain peptide substrates abolishes the potential contribution of other portions of GFI1 to LSD1 demethylase activity. We find substitution at the K239 SUMO acceptor residue impairs LSD1 binding to GFI1, suggesting additional elements outside the SNAG domain may regulate LSD1 binding and demethylase activity (Figure 2.8). Using full length GFI1 protein substrates may also reveal activities of LSD1 in the SNAG domain masked by our current approaches.

Characterizing GFI1 methylation status *in vivo*. It will also be informative to define GFI1 methylation status *in vivo*. SMYD2, SETD7, G9a and LSD1 are expressed in the majority of normal human tissues and are often misexpressed in cancer, suggesting complex GFI1 methylation statuses may exist *in vivo* (5). We have begun to characterize GFI1 methylation status in lymphoid leukemia, using

the T-lymphoblast cell line CCRF-CEM. Within GFI1 immunepurified material, we have detected GFI1 peptide fragments by mass spectrometry. We continue to optimize this method, including using an enzyme that will cut appropriately to allow detection of SNAG domain lysine methylation. *In silico* predictions suggest chymotrypsin digestion would appropriately produce a peptide fragment (^{6}LVKSKKAHSY^{15}) spanning the KSKK motif. We are currently proceeding with these technical considerations.

Developing antibody reagents capable of recognizing specific methylated forms of GFI1 will also be powerful to define GFI1 methylation status *in vivo*. Determining pertinent methylation states *in vitro* may also minimize the number of antibodies needed to be produced, minimizing the associated costs. Antibody reagents would allow for functional characterization of methylated GFI1 forms. Methyl specific antibodies could be applied in coimmunoprecipitation and ChIP-seq experiments to define methylation dependent GFI1 interactions as well as methylation dependent DNA binding in the genome.

Is there functional cross talk between GFI1 SUMOylation
and methylation?

Our findings that both SUMOylation and methylation contribute to GFI1 function through augmenting LSD1/CoREST binding raises the question of whether these modifications are coordinately regulated. We show here that SUMOylation machinery component PIAS3 and the methyltransferase SMYD2 both physically interact with GFI1 (Figure 2.1 and Figure 3.2). However, a direct

interaction between SMYD2 and PIAS3 has not been described, suggesting methylation and SUMOylation inputs on GFI1 may be independent of one another (6). SUMOylation/ubiquitination as well as methylation/acetylation crosstalk is prevalent in cells (7, 8). However, evidence suggesting SUMOylation/methylation crosstalk is currently limited (7, 8). Developing SUMO and methyl specific GFI1 antibodies will be useful for determining if and how these posttranslational modifications are regulated on GFI1, including the order of their addition and if the presence of one modification promotes or deters addition of the other.

Are methylation events within the GFI1 and p53 KSKK

motifs cross regulated?

We find the methyltransferases SMYD2, SETD7 and G9a have conserved activities for lysines within the GFI1 and p53 KSKK motifs. However, it remains unknown whether GFI1 and p53 methylation status affect one another, or whether there are functional consequences of such regulation.

Gfi1 antagonizes p53 apoptosis functions in multiple lymphoid leukemia mouse models (9). In leukemia cells, Gfi1 directly binds p53 and prevents expression of p53 responsive genes such as BAX (9). Gfi1 ablation also sensitizes mouse thymocytes to apoptosis following γ-irradiation (9). Furthermore, biallelic deletion of Gfi1 confers 100% survival in mice with Notch1 driven T-cell leukemia (9). These results implicate a direct functional relationship between GFI1 and p53 and GFI1 inhibition may represent a therapeutic opportunity in lymphoid leukemia.

In addition to survival benefits, biallelic Gfi1 deletion also increases K372

methylation on p53, a mark associated with p53 activation (9). These results suggest GFI1 may directly or indirectly influence p53 methylation status. However, the methylation status of GFI1 in lymphoid leukemia is currently unknown. It is also not known whether simply the presence of GFI1 dictates p53 methylation status or whether GFI1 methylation specifically augments p53 methylation.

GFI1 and p53 methyl specific antibodies would allow for characterization of GFI1/p53 methylation status cross talk, as well as the functional impact of this cross talk. These reagents would allow for determining the methylation status of GFI1 and p53 when they are physically associated. Antibody reagents would also be capable of determining the methylation status of each protein when it is bound at mutually exclusive and/or shared target genes. Correlating occupancy of methyl specific GFI1 and p53 forms with proximal histone marks and transcriptional status would also provide powerful insights into GFI1/p53 methylation dynamics and their methylation dependent functions in the genome.

Genome editing CRISPR-Cas9 technology will also be a powerful technique for characterizing GFI1 and p53 methylation crosstalk. CRISPR-Cas9 mediated substitution at methyl acceptor lysine residues would reveal how loss of methylation at one lysine affects methylation status at other lysines, both within the same protein or within the other respective protein.

Notch is a potent driver of T-cell leukemia and Gfi1 ablation is protective against Notch1 driven leukemias, likely through derepression of p53 function (9, 10). We have also demonstrated the Notch1-intracellular domain (N1-ICD) interacts directly with GFI1 (Figure 4.3) (11). Furthermore, N1ICD stabilizes LSD1

interaction with GFI1 (Figure 4.3). Future experiments will be needed to determine how Notch impacts GFI1 and p53 methylation statuses.

What is the functional breadth of transcription factor posttranslational modification in the human proteome?

Our findings here add GFI1 to a growing list of transcription factors and other nonhistone proteins regulated by lysine methylation. Transcription factors regulated by lysine methylation are predominantly factors with functions in cell fate, differentiation, and survival, many of which are disregulated in cancer. These master regulators include: p53, RB, E2F1, STAT3, MYOD, RELA/p65, AR, ERα, and CEBPβ (12, 13). The functional impact for many of these methylation events remains unclear.

Complete characterization of posttranslational modifications within the human transcription factor proteome remains a technically challenging effort. However, recent efforts have successfully profiled the monomethylation proteome in four cancer cell lines and a human liver cancer sample (14). This report indicates methylation events are highly context and cell type specific, with only 27 conserved methylation events observed across cell lines and primary tumor (14).

These methods are promising for discovering novel methylation events but are incapable of identifying methyltransferases and demethylases responsible for these events and are also unable to identify the functional consequences of these methylation events. Array peptide screening methods are promising for identifying

how specific posttranslational modifications alter function through augmenting protein/protein interactions. These methods use an array format to subject peptides with specific posttranslational modifications to protein bacterial expression libraries (15).

Array peptide screening methods could prove powerful for assessing the functional impact of specific posttranslational modifications, however; hypothesis driven investigation may be equally if not more productive. Hypothesis driven investigation enabled us to identify a shared KSKK motif between GFI1 and p53. We were subsequently able to demonstrate lysine methylation within this motif is conserved across both proteins. Furthermore, these hypothesis driven insights provided compelling questions for future study, more so than descriptive proteomic screening approaches.

Many sites of methylation in transcription factors and other nonhistone proteins reside within a loosely conserved KSKK motif (Table 1.1). As mentioned previously, GFI1 and p53 share a completely conserved KSKK motif. However, E2F1 is methylated within a KKSK motif. RELA/p65 is methylated within KSIMKK motif. RB is methylated within a KPLKK motif. TAF7 and TAF10 are both methylated within a KSKD motif. ERα is methylated within a KRSKK motif. IRF1 is methylated within a KSKS motif. NR1H4 is methylated within a KSKR motif (13). A Basic Local Alignment Search Tool (BLAST) search estimates roughly 330 genes encode a KSKK motif, 266 genes encode a KSKS motif, 195 genes encode a KSKR motif, 170 genes encode a KSKD motif, 20 genes encode a KRSKK motif, 17 genes encode a KPLKK, motif and 5 genes encode a KSIMKK motif (16). This

impressive number of proteins suggests many lysine methylation events within loosely conserved KSKK motifs remain to be identified as well as functionally characterized.

Conclusions: GFI1 is a model transcription factor for studying mechanisms of transcriptional regulation

Transcription factors are master regulators of cell fate, function, and survival. As such, understanding how transcription factors exert their transcriptional regulatory functions is critical for understanding their contribution to normal and disease states. During my thesis work, I studied how the transcription factor GFI1 carries out transcriptional repression through the action of its coeffector LSD1. My thesis work describes novel mechanisms of transcriptional repression by GFI1, and provides a rare mechanistic understanding for how posttranslational modifications modulate transcription factor functions. From the data generated during my thesis work, we propose GFI1 is subject to dual regulatory inputs where linker region SUMOylation and SNAG domain methylation serve to stabilize and activate LSD1/CoREST for H3K4 demethylation and repression of GFI1 target genes. We also speculate towards cross talk between GFI1 SNAG domain and p53 methylation status, given a shared set of methylation regulatory enzymes. Similar approaches to what we have taken here will allow future work to further define the functional breadth of transcription factor posttranslational modifications in normal and disease states.

References

1. **Moroy T, Vassen L, Wilkes B, Khandanpour C.** 2015. From cytopenia to leukemia: the role of Gfi1 and Gfi1b in blood formation. Blood **126:**2561.

2. **van der Meer LT, Jansen JH, van der Reijden BA.** 2010. Gfi1 and Gfi1b: key regulators of hematopoiesis. Leukemia **24:**1834–1843.

3. **Grimes HL, Chan TO, Zweidler-McKay PA, Tong B, Tsichlis PN.** 1996. The Gfi-1 proto-oncoprotein contains a novel transcriptional repressor domain, SNAG, and inhibits G1 arrest induced by interleukin-2 withdrawal. Mol Cell Bio **16:**6263.

4. **Person RE, Li F-Q, Duan Z, Benson KF, Wechsler J, Papadaki HA, Eliopoulos G, Kaufman C, Bertolone SJ, Nakamoto B, Papayannopoulou T, Grimes HL, Horwitz M.** 2003. Mutations in proto-oncogene GFI1 cause human neutropenia and target ELA2. Nat Genet **34:**308–312.

5. **Uhlén M, Westberg J, Wester K, Björling E, Wrethagen U, Agaton C, Xu LL, Szigyarto CA-K, Hober S, Amini B, Andersen E, Pontén F, Andersson A-C, Angelidou P, Asplund A, Asplund C, Berglund L, Bergström K, Brumer H, Cerjan D, Ekström M, Elobeid A, Eriksson C, Fagerberg L, Falk R, Fall J, Forsberg M, Björklund MG, Gumbel K, Halimi A, Hallin I, Hamsten C, Hansson M, Hedhammar M, Hercules G, Kampf C, Larsson K, Lindskog M, Lodewyckx W, Lund J, Lundeberg J, Magnusson K, Malm E, Nilsson P, Odling J, Oksvold P, Olsson I, Oster E, Ottosson J, Paavilainen L, Persson A, Rimini R, Rockberg J, Runeson M, Sivertsson A, Sköllermo A, Steen J, Stenvall M, Sterky F, Strömberg S, Sundberg M, Tegel H, Tourle S, Wahlund E, Waldén A, Wan J, Wernérus H, Westberg J, Wester K, Wrethagen U, Xu LL, Hober S, Pontén F.** 2005. A human protein atlas for normal and cancer tissues based on antibody proteomics. Mol Cell Proteomics **4:**1920–1932.

6. **Stark C, Breitkreutz B-J, Reguly T, Boucher L, Breitkreutz A, Tyers M.** 2005. BioGRID: a general repository for interaction datasets. Nucleic Acids Res **34:**D535–9.

7. **Latham JA, Dent SYR.** 2007. Cross-regulation of histone modifications. Nat Struct Mol Biol **14:**1017–1024.

8. **Hendriks IA, D'Souza RCJ, Yang B, Vries MV-D, Mann M, Vertegaal ACO.** 2014. Uncovering global SUMOylation signaling networks in a site-specific manner. Nat Struct Mol Biol **21:**927–936.

9. **Khandanpour C, Phelan JD, Vassen L, Schütte J, Chen R, Horman SR, Gaudreau M-C, Krongold J, Zhu J, Paul WE, Dührsen U, Göttgens B, Grimes HL, Möröy T.** 2013. Growth factor independence 1 antagonizes a p53-induced DNA damage response pathway in lymphoblastic leukemia. Cancer Cell **23:**200–214.

10. **Khandanpour C, Möröy T.** 2013. Gfi1 as a regulator of p53 and a therapeutic target for ALL. Oncotarget **4:**374.

11. **Yatim A, Benne C, Sobhian B, Laurent-Chabalier S, Deas O, Judde J-G, Lelievre J-D, Levy Y, Benkirane M.** 2012. NOTCH1 nuclear interactome reveals key regulators of its transcriptional activity and oncogenic function. Molecular Cell **48:**445–458.

12. **Biggar KK, Li SS-C.** 2014. Non-histone protein methylation as a regulator of cellular signalling and function. Nat Rev Mol Cell Biol **16:**5–17.

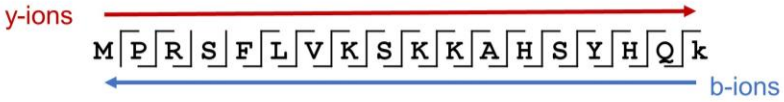
13. **Stark GR, Wang Y, Lu T.** 2010. Lysine methylation of promoter-bound transcription factors and relevance to cancer. Cell Research **21:**375–380.

14. **Wu Z, Cheng Z, Sun M, Wan X, Liu P, He T, Tan M, Zhao Y.** 2014. A chemical proteomics approach for global analysis of lysine monomethylome profiling. Mol Cell Proteomics **14:**329–339.

15. **Pless O, Kowenz-Leutz E, Dittmar G, Leutz A.** 2011. A differential proteome screening system for post-translational modification-dependent transcription factor interactions. Nat Protoc **6:**359–364.

16. **Altschul SF, Gish W, Miller W, Myers EW, Lipman DJ.** 1990. Basic local alignment search tool. J Mol Biol **215:**403–410.

Figure 4.1. SETD7 methylates K10 of the SNAG domain and G9a methylates K11 of the SNAG domain. Y and B ion fragmentation schematic is shown on top with each cleavage and resulting peptide indicated as staggered lines. Atomic mass units (AMU) of each y and b ion fragmentation species is shown in the table. A gain of 14.02 AMU, consistent with the addition of a single methyl group, at position 10 is observed in the presence of SETD7. Likewise, a 14.02 AMU gain at position 11 is observed in the presence of G9a.



#	Immonium	b	b-H2O	b-NH3	b (2+)	Seq	y	y-H2O	y-NH3	y (2+)	#
1	104.05	132.05	114.04	115.02	66.52	M					18
2	70.07	229.08	211.09	212.07	115.05	P	2267.22	2249.21	2250.20	1134.11	17
3	129.10	385.20	367.19	368.18	193.10	R	2170.17	2152.16	2153.14	1085.55	16
4	60.04	472.23	454.22	455.21	236.62	S	2014.07	1996.06	1997.04	1007.53	15
5	120.08	619.30	601.29	602.28	310.16	F	1927.04	1909.03	1910.01	964.02	14
6	86.10	732.39	714.38	715.36	366.69	L	1779.97	1761.96	1762.94	890.48	13
7	72.08	831.45	813.44	814.43	416.23	V	1666.88	1648.87	1649.86	833.94	12
8	101.07	959.55	941.54	942.52	480.28	K	1567.81	1549.79	1550.80	784.41	11
9	60.04	1046.58	1028.57	1029.56	523.79	S	1439.72	1421.70	1422.69	720.36	10
10	101.11	1174.68	1156.67	1157.65	587.84	K	1352.69	1334.68	1335.66	676.84	9
11	101.11	1302.78	1284.76	1285.74	651.89	K	1224.59	1206.57	1207.57	612.80	8
12	44.05	1373.82	1355.80	1356.78	687.40	A	1096.50	1078.49	1079.48	548.75	7
13	110.07	1510.87	1492.86	1493.84	755.94	H	1025.46	1007.46	1008.44	513.23	6
14	60.04	1597.90	1579.89	1580.87	799.45	S	888.40	870.38	871.37	444.70	5
15	136.08	1760.96	1742.95	1743.94	880.98	Y	801.37	783.36	784.34	401.19	4
16	110.07	1898.03	1880.01	1881.00	949.52	H	638.31	620.30	621.28	319.65	3
17	101.07	2026.08	2008.07	2009.05	1013.54	Q	501.25	483.24	484.22	251.12	2
18	327.19					K(+226.08)	373.19	355.18	356.16	187.10	1

Unmodified
NO MTase added

#	Immonium	b	b-H2O	b-NH3	b (2+)	Seq	y	y-H2O	y-NH3	y (2+)	#
1	104.05	132.05	114.04	115.02	66.52	M					18
2	70.07	229.10	211.09	212.07	115.05	P	2281.25	2263.23	2264.23	1141.12	17
3	129.10	385.20	367.19	368.18	193.10	R	2184.19	2166.18	2167.16	1092.59	16
4	60.04	472.24	454.22	455.21	236.62	S	2028.09	2010.07	2011.06	1014.54	15
5	120.08	619.30	601.29	602.28	310.16	F	1941.05	1923.04	1924.03	971.03	14
6	86.10	732.39	714.38	715.36	366.69	L	1794.00	1775.98	1776.96	897.49	13
7	72.08	831.46	813.44	814.43	416.23	V	1680.91	1662.89	1663.87	840.95	12
8	101.07	959.55	941.54	942.52	480.28	K	1581.84	1563.82	1564.81	791.42	11
9	60.04	1046.59	1028.57	1029.57	523.79	S	1453.74	1435.73	1436.73	727.37	10
10	115.12	1188.69	1170.68	1171.67	594.85	K(+14.02)	1366.70	1348.69	1349.68	683.85	9
11	101.11	1316.80	1298.78	1299.76	658.89	K	1224.60	1206.58	1207.58	612.80	8
12	44.05	1387.84	1369.82	1370.80	694.44	A	1096.51	1078.49	1079.48	548.75	7
13	110.07	1524.89	1506.87	1507.86	762.94	H	1025.47	1007.46	1008.44	513.23	6
14	60.04	1611.92	1593.91	1594.89	806.46	S	888.41	870.39	871.38	444.70	5
15	136.08	1774.98	1756.97	1757.95	887.99	Y	801.38	783.36	784.34	401.19	4
16	110.07	1912.05	1894.02	1895.01	956.52	H	638.33	620.30	621.28	319.65	3
17	101.07	2040.10	2022.09	2023.07	1020.55	Q	501.25	483.24	484.22	251.12	2
18	327.19					K(+226.08)	373.19	355.18	356.16	187.10	1

K10-methyl
+ SETD7

#	Immonium	b	b-H2O	b-NH3	b (2+)	Seq	y	y-H2O	y-NH3	y (2+)	#
1	104.05	132.05	114.04	115.02	66.52	M					18
2	70.07	229.08	211.09	212.07	115.05	P	2267.22	2249.21	2250.20	1134.11	17
3	129.10	385.20	367.19	368.18	193.10	R	2170.17	2152.16	2153.14	1085.55	16
4	60.04	472.23	454.22	455.21	236.62	S	2014.07	1996.06	1997.04	1007.53	15
5	120.08	619.30	601.29	602.28	310.16	F	1927.04	1909.03	1910.01	964.02	14
6	86.10	732.39	714.38	715.36	366.69	L	1779.97	1761.96	1762.94	890.48	13
7	72.08	831.45	813.44	814.43	416.23	V	1666.88	1648.87	1649.86	833.94	12
8	101.07	959.55	941.54	942.52	480.28	K	1567.81	1549.79	1550.80	784.41	11
9	60.04	1046.58	1028.57	1029.56	523.79	S	1439.72	1421.70	1422.69	720.36	10
10	101.11	1174.68	1156.67	1157.65	587.84	K	1352.69	1334.68	1335.66	676.84	9
11	101.11	1302.78	1284.76	1285.74	651.89	K	1224.59	1206.57	1207.57	612.80	8
12	44.05	1373.82	1355.80	1356.78	687.40	A	1096.50	1078.49	1079.48	548.75	7
13	110.07	1510.87	1492.86	1493.84	755.94	H	1025.46	1007.46	1008.44	513.23	6
14	60.04	1597.90	1579.89	1580.87	799.45	S	888.40	870.38	871.37	444.70	5
15	136.08	1760.96	1742.95	1743.94	880.98	Y	801.37	783.36	784.34	401.19	4
16	110.07	1898.03	1880.01	1881.00	949.52	H	638.31	620.30	621.28	319.65	3
17	101.07	2026.08	2008.07	2009.05	1013.54	Q	501.25	483.24	484.22	251.12	2
18	327.19					K(+226.08)	373.19	355.18	356.16	187.10	1

Unmodified
NO MTase added

#	Immonium	b	b-H2O	b-NH3	b (2+)	Seq	y	y-H2O	y-NH3	y (2+)	#
1	104.05	132.05	114.04	115.02	66.52	M					18
2	70.07	229.10	211.09	212.07	115.05	P	2281.25	2263.23	2264.23	1141.12	17
3	129.10	385.20	367.19	368.18	193.10	R	2184.19	2166.18	2167.16	1092.59	16
4	60.04	472.24	454.22	455.21	236.62	S	2028.09	2010.07	2011.06	1014.55	15
5	120.08	619.31	601.29	602.28	310.16	F	1941.05	1923.04	1924.03	971.03	14
6	86.10	732.39	714.38	715.36	366.69	L	1793.98	1775.97	1776.96	897.49	13
7	72.08	831.46	813.44	814.43	416.23	V	1680.91	1662.89	1663.87	840.95	12
8	101.07	959.56	941.54	942.52	480.28	K	1581.84	1563.82	1564.82	791.42	11
9	60.04	1046.60	1028.58	1029.57	523.79	S	1453.75	1435.73	1436.72	727.37	10
10	101.11	1174.69	1156.67	1157.65	587.84	K	1366.71	1348.69	1349.68	683.85	9
11	115.12	1316.79	1298.78	1299.76	658.89	K(+14.02)	1238.62	1220.60	1221.59	619.80	8
12	44.05	1387.84	1369.81	1370.80	694.44	A	1096.51	1078.49	1079.48	548.75	7
13	110.07	1524.89	1506.87	1507.86	762.94	H	1025.47	1007.45	1008.45	513.23	6
14	60.04	1611.92	1593.91	1594.89	806.46	S	888.41	870.39	871.38	444.70	5
15	136.08	1774.99	1756.98	1757.95	887.99	Y	801.38	783.36	784.34	401.19	4
16	110.07	1912.05	1894.03	1895.01	956.52	H	638.32	620.30	621.28	319.65	3
17	101.07	2040.10	2022.09	2023.07	1020.55	Q	501.25	483.24	484.22	251.12	2
18	327.19					K(+226.08)	373.19	355.18	356.16	187.10	1

K11-methyl
+ G9a

Figure 4.2. LSD1 does not demethylate K8 dimethylated SNAG domain peptide *in vitro*. Recombinant human LSD1 was incubated with either K8 dimethylated SNAG domain peptide or K4 dimethylated histone H3 peptide as indicated. LSD1 activity was assessed using the LSD1 Inhibitor Screening Assay Kit from Cayman Chemical according to manufacturer protocols. Enzyme activity was determined in a kinetic assay and normalized to a no LSD1 control.

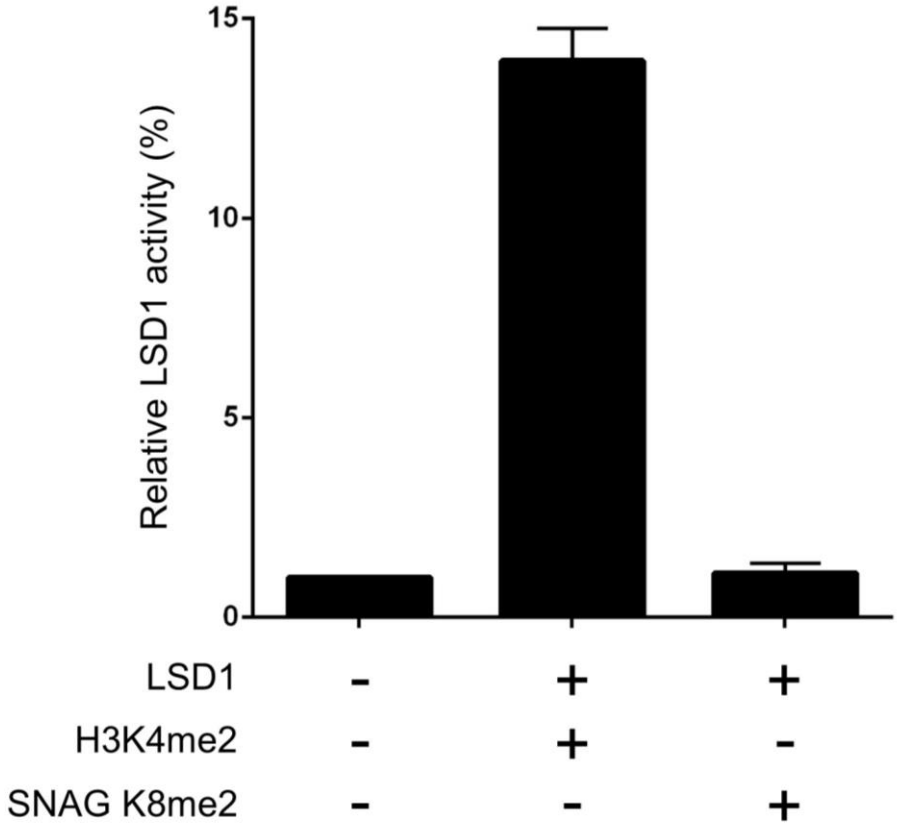
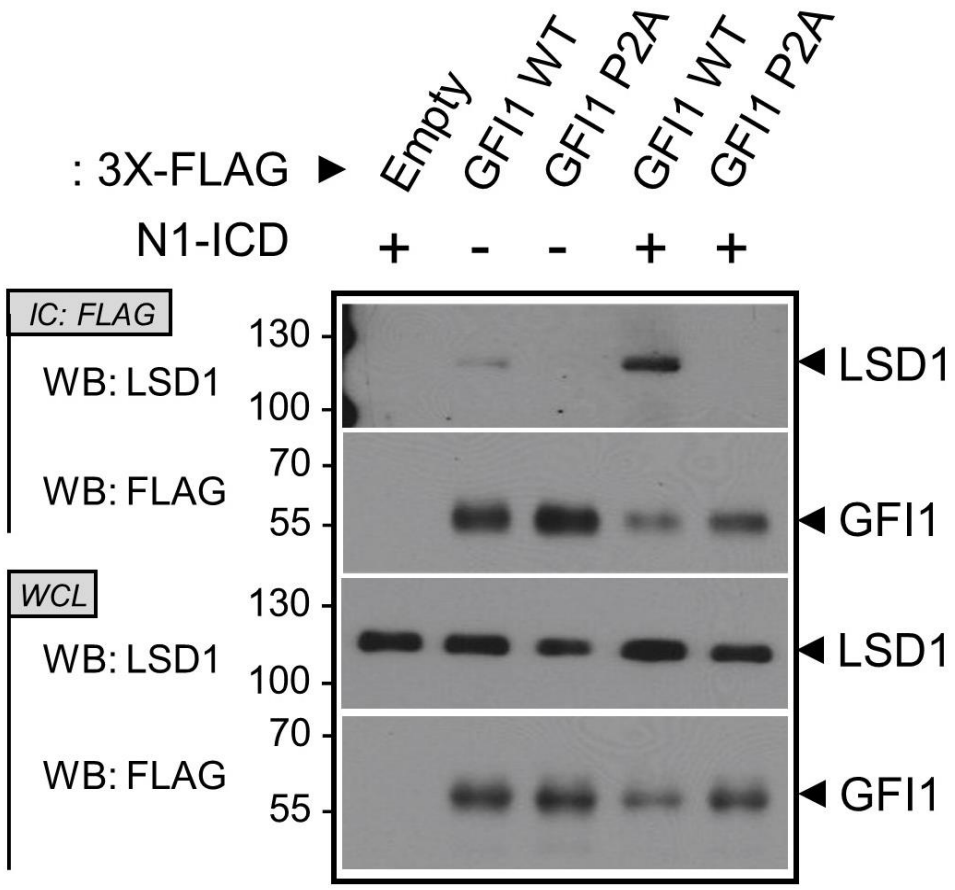

15
10
5
0
Relative LSD1 activity (%)
LSD1 - + +
H3K4me2 - + -
SNAG K8me2 - - +

Figure 4.3. Notch1 intracellular domain (N1ICD) stabilizes LSD1 interaction with GFI1. FLAG-tagged GFI1 forms were transiently expressed in Cos7L cells. GFI1 forms were immune purified with anti-FLAG antibody and coprecipitating endogenously expressed LSD1 was detected in immune complexes (IC) by western blot (WB).


: 3X-FLAG ►
Empty
GFI1 WT
GFI1 P2A
GFI1 WT
GFI1 P2A
N1-ICD
+ - - + +
IC: FLAG
WB: LSD1
WB: FLAG
WCL
WB: LSD1
WB: FLAG
130
100
70
55
130
100
70
55
LSD1
GFI1
LSD1
GFI1